

Oligomer search for  
length of 6

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: August 2, 2006, 09:10:59 ; Search time 27446 Seconds  
(without alignments)  
17665.574 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7582  
Sequence: 1 caacaatcggtataaacc.....tattaaatctgcactgtg 7582

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 6366136 seqs, 31973710525 residues

Word size : 6

Total number of hits satisfying chosen parameters: 12690550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : GenEmbl.\*

- 1: gb\_env.\*
- 2: gb\_pat.\*
- 3: gb\_ph.\*
- 4: gb\_pl.\*
- 5: gb\_pr.\*
- 6: gb\_ro.\*
- 7: gb\_sts.\*
- 8: gb\_sy.\*
- 9: gb\_un.\*
- 10: gb\_vi.\*
- 11: gb\_ov.\*
- 12: gb\_htg.\*
- 13: gb\_in.\*
- 14: gb\_om.\*
- 15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7582	100.0	7582	2	BD196245 Endogenet
2	7582	100.0	7582	2	BD267487 Endogenet
3	7582	100.0	7582	2	AX000966 Sequence
4	7582	100.0	7582	2	AX027480 Sequence
5	888	11.7	2372	2	BD196244 Endogenet
6	888	11.7	2372	2	BD267486 Endogenet
7	888	11.7	2372	2	AX000965 Sequence
8	888	11.7	2372	2	AX027479 Sequence
9	888	11.7	2372	5	AF072504 Homo sapi
10	779	10.3	2599	2	BD221806 Nucleic s
11	779	10.3	2599	2	AR699755 Sequence
12	779	10.3	8523	2	AX007978 Sequence
13	779	10.3	8523	2	AR612522 Sequence
14	779	10.3	10222	5	AY101582 Homo sapi
15	779	10.3	10222	5	AY101582 Homo sapi
16	779	10.3	10499	2	BD221808 Nucleic s
17	779	10.3	10499	2	AR699757 Sequence
18	779	10.3	10499	2	AX007980 Sequence

19	779	10.3	5693	2	AX329572 Sequence
20	779	10.3	5693	5	HSAC000064 Human BAC
21	779	10.3	149194	5	AC007566 Homo sapi
22	731	9.6	1617	5	AF156963 Homo sapi
23	731	9.6	1948	2	BD196239 Endogenet
24	731	9.6	1948	2	BD267481 Endogenet
25	731	9.6	1948	2	AX000960 Sequence
26	731	9.6	1948	2	AX027474 Sequence
27	731	9.6	1948	5	AF072505 Homo sapi
28	731	9.6	2055	2	BD221827 Nucleic s
29	731	9.6	2055	2	AR699776 Sequence
30	731	9.6	2055	2	AX007999 Sequence
31	731	9.6	2694	5	AF520483S2 Homo sapi
32	731	9.6	2694	5	AF520485S2 Homo sapi
33	731	9.6	2694	5	AF520493S2 Homo sapi
34	731	9.6	2694	5	AF520495S2 Homo sapi
35	731	9.6	2694	5	AF520496 Homo sapi
36	731	9.6	2694	5	AF520497S2 Homo sapi
37	731	9.6	2694	5	AF520498 Homo sapi
38	731	9.6	2694	5	AF520500 Homo sapi
39	731	9.6	2694	5	AF520501S2 Homo sapi
40	731	9.6	2694	5	AF520502 Homo sapi
41	731	9.6	2694	5	AF520503S2 Homo sapi
42	731	9.6	2694	5	AF520509S2 Homo sapi
43	731	9.6	2694	5	AF520511S2 Homo sapi
44	731	9.6	2694	5	AF520513S2 Homo sapi
45	731	9.6	2694	5	AF520515S2 Homo sapi
46	731	9.6	2694	5	AF520523S2 Homo sapi
47	731	9.6	2694	5	AF520525S2 Homo sapi
48	731	9.6	2694	5	AF520527S2 Homo sapi
49	731	9.6	2694	5	AF520528 Homo sapi
50	731	9.6	2694	5	AF520529S2 Homo sapi
51	731	9.6	2694	5	AF520531S2 Homo sapi
52	731	9.6	2694	5	AF520533S2 Homo sapi
53	731	9.6	2694	5	AF520535S2 Homo sapi
54	731	9.6	2694	5	AF520537S2 Homo sapi
55	731	9.6	2694	5	AF520539S2 Homo sapi
56	731	9.6	2694	5	AF520541S2 Homo sapi
57	731	9.6	2694	5	AF520542 Homo sapi
58	731	9.6	2694	5	AF520543S2 Homo sapi
59	731	9.6	2694	5	AF520544 Homo sapi
60	731	9.6	2694	5	AF520545S2 Homo sapi
61	731	9.6	2694	5	AF520546 Homo sapi
62	731	9.6	2782	2	BD196241 Endogenet
63	731	9.6	2782	2	BD267483 Endogenet
64	731	9.6	2782	2	AX000962 Sequence
65	731	9.6	2782	2	AX027476 Sequence
66	728	9.6	2782	2	AX133396 Sequence
67	728	9.6	2930	2	AX355872 Sequence
68	728	9.6	2930	5	AF208161 Homo sapi
69	728	9.6	10222	5	AY101583 Homo sapi
70	724	9.5	2938	2	BD196236 Endogenet
71	724	9.5	2938	2	BD267478 Endogenet
72	724	9.5	2938	2	AX000957 Sequence
73	724	9.5	2938	2	AX027471 Sequence
74	680	9.0	1624	5	AF072499 Homo sapi
75	680	9.0	1860	5	AF506835 Homo sapi
76	680	9.0	2694	5	AF513360 Homo sapi
77	680	9.0	2694	5	AF520477S2 Homo sapi
78	680	9.0	2694	5	AF520479S2 Homo sapi
79	680	9.0	2694	5	AF520481S2 Homo sapi
80	680	9.0	2694	5	AF520487S2 Homo sapi
81	680	9.0	2694	5	AF520489S2 Homo sapi
82	680	9.0	2694	5	AF520505S2 Homo sapi
83	680	9.0	2694	5	AF520517S2 Homo sapi
84	680	9.0	2694	5	AF520519S2 Homo sapi
85	680	9.0	2694	5	AF520521S2 Homo sapi
86	680	9.0	2694	5	AF520529S2 Homo sapi
87	680	9.0	2694	5	AF520545S2 Homo sapi
88	680	9.0	2694	5	AF520559S2 Homo sapi
89	680	9.0	2694	5	AF520561S2 Homo sapi
90	680	9.0	2946	2	ARI77269 Sequence
			2946	2	BD086133 Secreted

## ALIGNMENTS

```
RESULT 1
BD196245
LOCUS          7582 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION    Endogenous retroviral sequences, associated with autoimmune
               diseases and/or with pregnancy disorders.
ACCESSION     BD196245
VERSION       BD196245.1 GI:33006015
KEYWORDS      JP 2002512530-A/11.
SOURCE        unidentified
ORGANISM      unidentified
               unclassified sequences.
REFERENCE     1 (bases 1 to 7582)
AUTHORS       Beseme, F., Blond, J.L., Bouton, O., Mandrand, B. and Mallet, F.
TITLE         Endogenous retroviral sequences, associated with autoimmune
               diseases and/or with pregnancy disorders
JOURNAL       Patent: JP 2002512530-A 11 23-APR-2002;
               BIO MERIEUX
COMMENT       OS Unidentified
               PN JP 2002512530-A/11
               PD 23-APR-2002
               PP 06-JUL-1998 JP 1999508244
               PR 07-JUL-1997 FR 97/08815
               PI FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD
               MANDRAND,
               PI FRANCOIS MALLET
               PC C12N15/48, C07K14/15, C12Q1/68, C07K16/10, G01N33/569 CC
               Strandedness: Single;
               CC Topology: Linear;
               CC Endogenous retroviral sequences, associated with autoimmune
               diseases
               CC and/or with pregnancy disorders
               PH Key Location/Qualifiers
               FT source 1..7582 /organism='Unidentified'.
               FEAT source
                 Location/Qualifiers
                 1..7582
                 /organism='unidentified'
                 /mol_type='genomic DNA'
                 /db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 7582; DB 2; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAAATCGGGATATAAACCAGGCGATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60
DB 1 CAACAAATCGGGATATAAACCAGGCGATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60

QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACCTATTAAATCTGCAACTGCA 120
DB 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACCTATTAAATCTGCAACTGCA 120

QY 121 CTTCTTCGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCAACCGTCCACCACTGC 180
DB 121 CTTCTTCGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCAACCGTCCACCACTGC 180

QY 181 TGTTCGCCACCCGCGANACCTGCGCTGACTCCCATCCCTCTGGAATCTCGAGGGGTGC 240
DB 181 TGTTCGCCACCCGCGANACCTGCGCTGACTCCCATCCCTCTGGAATCTCGAGGGGTGC 240

QY 241 CGCTGTGCTCTGATCCAGCGCGCCCATTTGCCCTCCCAATTCGGGCTAAAGGCTTGC 300
DB 241 CGCTGTGCTCTGATCCAGCGCGCCCATTTGCCCTCCCAATTCGGGCTAAAGGCTTGC 300

QY 301 CATTTGNCCTGCAAGGCTGAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTTACT 360
DB 301 CATTTGNCCTGCAAGGCTGAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTTACT 360
```

```
QY 361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATTAACACTTACCACA 420
DB 361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATTAACACTTACCACA 420

QY 421 TGGCCCAAGATTCCATTCCTTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAAGAAATAC 480
DB 421 TGGCCCAAGATTCCATTCCTTTGGAATCCGTGAGGSCAACGAACTCCAGGTCAAGAAATAC 480

QY 481 GARGCTTCCACCATCTTTGGAAGCGGCTGCTACCTCTTGGAAAGTGGTTTCAACCACTC 540
DB 481 GARGCTTCCACCATCTTTGGAAGCGGCTGCTACCTCTTGGAAAGTGGTTTCAACCACTC 540

QY 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGCRAACMSRACGACATCC 600
DB 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGCRAACMSRACGACATCC 600

QY 601 MAAGTGATGGGAAACGTTTCCCGCAAGACAAAAACGCCCTTAAGACGTTATTTCTGGARAAT 660
DB 601 MAAGTGATGGGAAACGTTTCCCGCAAGACAAAAACGCCCTTAAGACGTTATTTCTGGARAAT 660

QY 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGCACTGCC 720
DB 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGCACTGCC 720

QY 721 GCCTGGCACTCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACVCTCTTTTG 780
DB 721 GCCTGGCACTCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACVCTCTTTTG 780

QY 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAAATTTCTTTTCAATTAAGACAAA 840
DB 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAAATTTCTTTTCAATTAAGACAAA 840

QY 841 CTCACAAATATGTAATAAAGTGTATTTATGCTTACAGAAAGCTTTCAGAGCTACCTCC 900
DB 841 CTCACAAATATGTAATAAAGTGTATTTATGCTTACAGAAAGCTTTCAGAGCTACCTCC 900

QY 901 CTATCCAGCATCCCGACTCTTCCCCAMTAAATAGGACCCCTTCAACCCAAATGG 960
DB 901 CTATCCAGCATCCCGACTCTTCCCCAMTAAATAGGACCCCTTCAACCCAAATGG 960

QY 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT 1020
DB 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT 1020

QY 1021 TATGACCCCTCCCAAGAGTGGGAGGAGAAAGAAATTCGGCCACAGCAGTGCATGCTT 1080
DB 1021 TATGACCCCTCCCAAGAGTGGGAGGAGAAATTCGGCCACAGCAGTGCATGCTT 1080

QY 1081 TTTTCTCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAATCCT 1140
DB 1081 TTTTCTCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAATCCT 1140

QY 1141 GATGCTATATTTGRTGTTTCAAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAT 1200
DB 1141 GATGCTATATTTGRTGTTTCAAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAT 1200

QY 1201 ATATATGCTACCTGCTAAATTCAGACACTAACCCCAATGAGAGAGTGCCACCAATAC 1260
DB 1201 ATATATGCTACCTGCTAAATTCAGACACTAACCCCAATGAGAGAGTGCCACCAATAC 1260

QY 1261 AGCCTGAGGTTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAA 1320
DB 1261 AGCCTGAGGTTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAA 1320

QY 1321 GAAGGAAAGANAATGATTTCCCAAGGCGCAGCAGTTCAGCTTASACCTCATTTG 1380
DB 1321 GAAGGAAAGANAATGATTTCCCAAGGCGCAGCAGTTCAGCTTASACCTCATTTG 1380

QY 1381 GGGACACAGAAATCAGTAACATGGGAGATTTGCTGAGACATTTGCTAACTTGTGTC 1440
DB 1381 GGGACACAGAAATCAGTAACATGGGAGATTTGCTGAGACATTTGCTAACTTGTGTC 1440

QY 1441 TASAAGGACTAAGGAAAACTASGAGAAATCTAYGAAATTAATCTCAATGATGTCCACATA 1500
```



1441 TASAAGGACTAAGGAAACTASGAAGAAATCTAYGAATTAATCTCAATGATGTCCACATA 1500  
1501 ACACAGGGAGGGAAGAAATCCTACTGCTTTCTGGAGAGACTAAGGAGGACATTGAG 1560  
1501 ACACAGGGAGGGAAGAAATCCTACTGCTTTCTGGAGAGACTAAGGAGGACATTGAG 1560  
1561 GAAGCGTGCTCTCTGTCACCTGACCTCTCTGAAGGCCCACTAATCTTAAGCGTAAGTT 1620  
1561 GAAGCGTGCTCTCTGTCACCTGACCTCTCTGAAGGCCCACTAATCTTAAGCGTAAGTT 1620  
1621 TATCACTCAGTCAGCTGCAGACATTAAGAAAAAATCAAAAGTCTCCGCTAGGCCCGGAG 1680  
1621 TATCACTCAGTCAGCTGCAGACATTAAGAAAAAATCAAAAGTCTCCGCTAGGCCCGGAG 1680  
1681 CAANAACCTTAGAACCCTATTGAATTTGGCAACCTCTGCTTTTATTAATAGAGATCAGGAG 1740  
1681 CAANAACCTTAGAACCCTATTGAATTTGGCAACCTCTGCTTTTATTAATAGAGATCAGGAG 1740  
1741 GAGCAGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCAACCGCTTTAGTCATGACCTT 1800  
1741 GAGCAGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCAACCGCTTTAGTCATGACCTT 1800  
1801 CAGGCAAGTGGAATTTGGAGGCTCTCGAAAAAGGAAAGCTGGGCAAAATTGAATGCTTAA 1860  
1801 CAGGCAAGTGGAATTTGGAGGCTCTCGAAAAAGGAAAGCTGGGCAAAATTGAATGCTTAA 1860  
1861 TAGGGCTTGGCTTCCAGTCGGGCTTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAG 1920  
1861 TAGGGCTTGGCTTCCAGTCGGGCTTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAG 1920  
1921 TAAGCGCCCTTCGTCATGCCCCCTTATTTCAAGGGAATCACTGAAGGCCCACTGCCCC 1980  
1921 TAAGCGCCCTTCGTCATGCCCCCTTATTTCAAGGGAATCACTGAAGGCCCACTGCCCC 1980  
1981 CAGGGGACAAAGGCTTTTGAAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGGACTG 2040  
1981 CAGGGGACAAAGGCTTTTGAAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGGACTG 2040  
2041 AGGGTGCTGGGGCAAGGCCATCCCATGCCATCACTCCCTCAGAGGCCCTGGGTATGCTT 2100  
2041 AGGGTGCTGGGGCAAGGCCATCCCATGCCATCACTCCCTCAGAGGCCCTGGGTATGCTT 2100  
2101 GACCATTAAGGGCCAGGAAGGTTGTCTCTGACACCTGTCGGTCTTTCTTGTAGCTTACT 2160  
2101 GACCATTAAGGGCCAGGAAGGTTGTCTCTGACACCTGTCGGTCTTTCTTGTAGCTTACT 2160  
2161 CTTCGTCCCGGACAACTGCTCTCCAGATCTCTCACTATCTGAGGGGTCCTTAAGAG 2220  
2161 CTTCGTCCCGGACAACTGCTCTCCAGATCTCTCACTATCTGAGGGGTCCTTAAGAG 2220  
2221 GGCAGTCACTAGATACTTTTCCAGCCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTT 2280  
2221 GGCAGTCACTAGATACTTTTCCAGCCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTT 2280  
2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTCTGTTAGGGAGAGACATTTCTA 2340  
2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTCTGTTAGGGAGAGACATTTCTA 2340  
2341 GCANAAGCAGGGCCATTAACCTGAACATAGAGAGGAACACCCGTTTGTGTNCC 2400  
2341 GCANAAGCAGGGCCATTAACCTGAACATAGAGAGGAACACCCGTTTGTGTNCC 2400  
2401 CTGCTTTAGGAGGAATTAATCTGAAGTCTGGGCAACAGAGGACCAATATGGAGGAGC 2460  
2401 CTGCTTTAGGAGGAATTAATCTGAAGTCTGGGCAACAGAGGACCAATATGGAGGAGC 2460  
2461 CAAGAATGCCCGTCTGTTCAAGTTAAATAAGGATTTCACTTCTCTTCCCTACCAAA 2520  
2461 CAAGAATGCCCGTCTGTTCAAGTTAAATAAGGATTTCACTTCTCTTCCCTACCAAA 2520  
2521 GGCAGTACCCCTCAGACCCAGGCCCAACAGAGATTTCCAAAGATTTGTTAAGGACTTAA 2580

2521 GGCAGTACCCCTCAGACCCAGGCCCAACAGAGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
2581 AAGCCCAAGGCTTAGTAAACCATGATTAACCTCCCTGAGTAAATTCCTGAGTGGATTGAG 2640  
2581 AAGCCCAAGGCTTAGTAAACCATGATTAACCTCCCTGAGTAAATTCCTGAGTGGATTGAG 2640  
2641 GAGGCACAGAAACCCAGTGGAAGTGGAGGTTAGTGCAGAGATCTCAGGATTTATCAATGG 2700  
2641 GAGGCACAGAAACCCAGTGGAAGTGGAGGTTAGTGCAGAGATCTCAGGATTTATCAATGG 2700  
2701 AGGCCGTTGTCTCTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTTCCCAATATC 2760  
2701 AGGCCGTTGTCTCTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTTCCCAATATC 2760  
2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGACCTTTCAGGATAGCTTCTTCTGATCCCTG 2820  
2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGACCTTTCAGGATAGCTTCTTCTGATCCCTG 2820  
2821 TACATCTCTGACTCTCAATTTCTTTGTTTGGCTTTGAAGATACTTTCAAAACCCATCTCAAC 2880  
2821 TACATCTCTGACTCTCAATTTCTTTGTTTGGCTTTGAAGATACTTTCAAAACCCATCTCAAC 2880  
2881 TCACCTCTGACTCTCAATTTCTTTGTTTGGCTTTGAAGATACTTTCAAAACCCATCTCAAC 2940  
2881 TCACCTCTGACTCTCAATTTCTTTGTTTGGCTTTGAAGATACTTTCAAAACCCATCTCAAC 2940  
2941 TAGCCCAAGACTTTGAGYCACTTTCATACCTGACACTCTTCTCTCTGTAAGTGGATG 3000  
2941 TAGCCCAAGACTTTGAGYCACTTTCATACCTGACACTCTTCTCTCTGTAAGTGGATG 3000  
3001 ATTTACTTTTTRGCGCTTTCAGAAACCTTTGTCGATCAAGCCACCCAAAGCRCTTTMA 3060  
3001 ATTTACTTTTTRGCGCTTTCAGAAACCTTTGTCGATCAAGCCACCCAAAGCRCTTTMA 3060  
3061 ATTTCTCTCGVACCTGTGCTTACAGGTTTCCAAACSAARACTCTCTCTCTCACAGC 3120  
3061 ATTTCTCTCGVACCTGTGCTTACAGGTTTCCAAACSAARACTCTCTCTCTCACAGC 3120  
3121 AGGTTAAATTAATCTTAGGCTTAARATTTATCCAAAGGCCACCAAGGCCCTCAGTGAGGAAYR 3180  
3121 AGGTTAAATTAATCTTAGGCTTAARATTTATCCAAAGGCCACCAAGGCCCTCAGTGAGGAAYR 3180  
3181 TCCAGCCTATCTGCTTATCTCTATCTCAATCAAAACCCCTAAAGCACTAAGGRRTTCTTTG 3240  
3181 TCCAGCCTATCTGCTTATCTCTATCTCAATCAAAACCCCTAAAGCACTAAGGRRTTCTTTG 3240  
3241 GCRTAAAYAGGTTCTGCGGAATWATGATTTCCAGGTTTGGCAATATAGCCAGGCTTTA 3300  
3241 GCRTAAAYAGGTTCTGCGGAATWATGATTTCCAGGTTTGGCAATATAGCCAGGCTTTA 3300  
3301 WATACASTAAATTAAGGAACTCAGAAAGCCCAATACCTTTTARTAGATGGAYAMCTGAA 3360  
3301 WATACASTAAATTAAGGAACTCAGAAAGCCCAATACCTTTTARTAGATGGAYAMCTGAA 3360  
3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCCAGTGTAAAGYT 3420  
3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCCAGTGTAAAGYT 3420  
3421 TGCCAAACRGGGCAAGACTTTTSTTATATCAAGAAAAAAGCAGAAAYAGCTCTRGA 3480  
3421 TGCCAAACRGGGCAAGACTTTTSTTATATCAAGAAAAAAGCAGAAAYAGCTCTRGA 3480  
3481 GTCTTACACAGRTCCRAGGAYGAGCTTGCATTTTAYGGGTAGTGGTGGCAGTACAGTAKTA 3540  
3481 GTCTTACACAGRTCCRAGGAYGAGCTTGCATTTTAYGGGTAGTGGTGGCAGTACAGTAKTA 3540  
3541 GATGTAGTGGCAAGGTTTTCATTTTAYGGGTAGTGGTGGCAGTACAGTAKTA 3600  
3541 GATGTAGTGGCAAGGTTTTCATTTTAYGGGTAGTGGTGGCAGTACAGTAKTA 3600  
3601 GTATCTGAAGCAGTTAAATTAATCAGGGAGAGATTTTCTGTGTGACATCTCATGAK 3660  
3601 GTATCTGAAGCAGTTAAATTAATCAGGGAGAGATTTTCTGTGTGACATCTCATGAK 3660

3661 GTGAATGCTACTCACTGCTAAAGGAGACTTGTGGCTGTGACAGCAACVGTCTTAA 3720  
Db |||||  
3661 GTGAATGCTACTCACTGCTAAAGGAGACTTGTGGCTGTGACAGCAACVGTCTTAA 3720  
Qy |||||  
3721 TTTCAAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGACATTTGTGCAACTCTTAAC 3780  
Db |||||  
3721 TTTCAAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGACATTTGTGCAACTCTTAAC 3780  
Qy |||||  
3781 CCAGTNCATTTCTTCCAGACAATGAAGAAAGATARAAYATACTGTCAACAARTAAAT 3840  
Db |||||  
3781 CCAGTNCATTTCTTCCAGACAATGAAGAAAGATARAAYATACTGTCAACAARTAAAT 3840  
Qy |||||  
3841 TCTCAAACTATGCCACTCGAGGACCTTYPAGARGTTCCVTTGACATGATCCYACCTT 3900  
Db |||||  
3841 TCTCAAACTATGCCACTCGAGGACCTTYPAGARGTTCCVTTGACATGATCCYACCTT 3900  
Qy |||||  
3901 CAACCTGTATCTGATGGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
Db |||||  
3901 CAACCTGTATCTGATGGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
Qy |||||  
3961 AGTGTCTAGTGAATATGGAATATTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGTCTYA 4020  
Db |||||  
3961 AGTGTCTAGTGAATATGGAATATTTGAAAGTAATCCCCTCACTCCAGGAACTAGTGTCTYA 4020  
Qy |||||  
4021 GCTRCAGAACTAATAGCCYTCAYTKGGCACTAGAAATTTAGGAGAGRAAAAGGGYAAA 4080  
Db |||||  
4021 GCTRCAGAACTAATAGCCYTCAYTKGGCACTAGAAATTTAGGAGAGRAAAAGGGYAAA 4080  
Qy |||||  
4081 TATATATACAGACTCTRTATATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
Db |||||  
4081 TATATATACAGACTCTRTATATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
Qy |||||  
4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db |||||  
4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy |||||  
4201 ATTATATYGGGWTACAGAAACCTTAGAGAGTGGAGTCTTACACTGCGGGTATCA 4260  
Db |||||  
4201 ATTATATYGGGWTACAGAAACCTTAGAGAGTGGAGTCTTACACTGCGGGTATCA 4260  
Qy |||||  
4261 NAAAGGAAGRAAGGGAATASAGRGAAATGCAAGCAKATATTAAGCMAAAAGAGC 4320  
Db |||||  
4261 NAAAGGAAGRAAGGGAATASAGRGAAATGCAAGCAKATATTAAGCMAAAAGAGC 4320  
Qy |||||  
4321 TGCAAGGAGGACCTCCATTAAGAAATGCTTATTAACCTTCCTTAGTATAGGGTAATCC 4380  
Db |||||  
4321 TGCAAGGAGGACCTCCATTAAGAAATGCTTATTAACCTTCCTTAGTATAGGGTAATCC 4380  
Qy |||||  
4381 CTTCCGGGNAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGNAACCTCAGG 4440  
Db |||||  
4381 CTTCCGGGNAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGNAACCTCAGG 4440  
Qy |||||  
4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
Db |||||  
4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
Qy |||||  
4501 TATCAATGGAATTACTTAAACCTTCATCAAACTTTCACTTAGGCATCGATAGCAC 4560  
Db |||||  
4501 TATCAATGGAATTACTTAAACCTTCATCAAACTTTCACTTAGGCATCGATAGCAC 4560  
Qy |||||  
4561 CCATCARATGGCAATCATTTATTTACTGGACAGGCTTTTCAAACTATCAAGCARAT 4620  
Db |||||  
4561 CCATCARATGGCAATCATTTATTTACTGGACAGGCTTTTCAAACTATCAAGCARAT 4620  
Qy |||||  
4621 AKTCAAGGCTGTGAATGTGCCARAAATAATCCCCTGCCCTYATCGCAAGCTCCTTC 4680  
Db |||||  
4621 AKTCAAGGCTGTGAATGTGCCARAAATAATCCCCTGCCCTYATCGCAAGCTCCTTC 4680  
Qy |||||  
4681 AGGAACAAARAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCAACAG 4740  
Db |||||  
4681 AGGAACAAARAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCAACAG 4740

4741 CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTTACCGGTTGGCA 4800  
Db |||||  
4741 CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTTACCGGTTGGCA 4800  
Qy |||||  
4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Db |||||  
4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Qy |||||  
4861 ATAAATTTCCAGATTTGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTTTCCAG 4920  
Db |||||  
4861 ATAAATTTCCAGATTTGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTTTCCAG 4920  
Qy |||||  
4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACCTCGGCC 4980  
Db |||||  
4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACCTCGGCC 4980  
Qy |||||  
4981 TGAAGGCAACAGTCTCAGGGAAGGTGAGAAAAATGAATGAATAAAYACTCAAGGAATCTA 5040  
Db |||||  
4981 TGAAGGCAACAGTCTCAGGGAAGGTGAGAAAAATGAATGAATAAAYACTCAAGGAATCTA 5040  
Qy |||||  
5041 AAAAAGCAACCCAGGAAACCCACTCACATGGCTGTCTGTGGCTTATAGCCTTAAA 5100  
Db |||||  
5041 AAAAAGCAACCCAGGAAACCCACTCACATGGCTGTCTGTGGCTTATAGCCTTAAA 5100  
Qy |||||  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Db |||||  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Qy |||||  
5161 CCCTTCAATAACCAATGACTTGTGTGACCCCAAGACAGCCAACTTAGTGTGAGACATCA 5220  
Db |||||  
5161 CCCTTCAATAACCAATGACTTGTGTGACCCCAAGACAGCCAACTTAGTGTGAGACATCA 5220  
Qy |||||  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAAACTTTAAGGAACCTTATCCCTGAGAG 5280  
Db |||||  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAAACTTTAAGGAACCTTATCCCTGAGAG 5280  
Qy |||||  
5281 AGGAAAAAGAACTATTCCACCCWMTGACATGTGTATTAGTCAAGTCCCTTCTCTAAT 5340  
Db |||||  
5281 AGGAAAAAGAACTATTCCACCCWMTGACATGTGTATTAGTCAAGTCCCTTCTCTAAT 5340  
Qy |||||  
5341 CCCATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATTTATCCCAACTG 5400  
Db |||||  
5341 CCCATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATTTATCCCAACTG 5400  
Qy |||||  
5401 CGTTTAAAGTGGCTGGAGTGTGGATACATCACTTGAGTCAAACTCTGGATAC 5460  
Db |||||  
5401 CGTTTAAAGTGGCTGGAGTGTGGATACATCACTTGAGTCAAACTCTGGATAC 5460  
Qy |||||  
5461 TGCCAAAGGAAACCTGAAATCCAGGAGCAACGCTAGTCTTCTGTGAAACCTCTAGAGG 5520  
Db |||||  
5461 TGCCAAAGGAAACCTGAAATCCAGGAGCAACGCTAGTCTTCTGTGAAACCTCTAGAGG 5520  
Qy |||||  
5521 ATTTGCGCTCTCTTTCAAAACCAACAGGAGGAAAGTAATAAATCATATAATCCCCC 5580  
Db |||||  
5521 ATTTGCGCTCTCTTTCAAAACCAACAGGAGGAAAGTAATAAATCATATAATCCCCC 5580  
Qy |||||  
5581 ATGSCCTCCCTTATCATATTTTCTKTASTGTSTTTTACCTCTTTTCACTCTCACT 5640  
Db |||||  
5581 ATGSCCTCCCTTATCATATTTTCTKTASTGTSTTTTACCTCTTTTCACTCTCACT 5640  
Qy |||||  
5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db |||||  
5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy |||||  
5701 ATGACGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCC 5760  
Db |||||  
5701 ATGACGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCC 5760  
Qy |||||  
5761 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTTGCACCTCTTTGCAATG 5820  
Db |||||  
5761 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTTGCACCTCTTTGCAATG 5820  
Qy |||||  
5821 CATGCAAACTACTCATTATTGGACAGGAAAAATGATTAATCTTGTCTCTGAGGACTT 5880

	Db	5821	 CATCAAAATACTCATTATTTGCACAGGAAAAAATGATTAATCTTAGTGTCTCGTGAGGACTT	5880
	Qy	5881	GGAGTCACCTGCTCTTGCGACTTACTTCACCACCAACTGGTATGCTGATGGGGTGGAGTT	5940
	Db	5881	GGAGTCACCTGCTCTGTTGGA CTTACTTCACCCCAAAC TGGTATGTCGATGGGGTGGAGTT	5940
	Qy	5941	CAAGATCAGGACAAGAGAAAAACATGTGTAAAAGAAGTAAATCTCCAACTCACCSGGGTACAT	6000
	Db	5941	CAAGATCAGGACAAGAGAAAAACATGTAAAAGAAGTAAATCTCCAACTCACCSGGGTACAT	6000
	Qy	6001	GGCACCTCTAGCCCCCTACAAGGACTAGATCTCTCAAACTACATGNAACCCCTCCGTACC	6060
	Db	6001	GGCACCTCTTAGCCCCCTACAAGGACTAGATCTCTCAAACTACATGNAACCCCTCCGTACC	6060
	Qy	6061	CATACTCGCCTGGTGAAGCCTATTTAATPACCACCCCTCACTGGGCTC CATGAGGTCTCGGCC	6120
	Db	6061	CATACTCGCCTGGTGAAGCCTATTTAATPACCACCCCTCACTGGGCTC CATGAGGTCTCGGCC	6120
	Qy	6121	CAAAACCCCTACTAACCTGTTGGATGATGCTCCCCTGAACTTCARGCCCATATGTTCAATC	6180
	Db	6121	CAAAACCCCTACTAACCTGTTGGATGATGCTCCCCTGAACTTCARGCCCATATGTTCAATC	6180
	Qy	6181	CTGTACTCTGAA CAATGGAACA ACTTCAGACACAGAAATAAACACCACTTCCTCCCAACAA	6240
	Db	6181	CTGTACTCTGAA CAATGGAACA ACTTCAGACACAGAAATAAACACCACTTCCTCCCAACAA	6240
	Qy	6241	GGACCTCTTGTTC CCAATSGGAAATAACCCATACCTCAAACTCTCACTGTGTAAATTT	6300
	Db	6241	GGACCTCTTGTTC CCAATSGGAAATAACCCATACCTCAAACTCTCACTGTGTAAATTT	6300
	Qy	6301	AGCAATACTACATACACAACAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAA	6360
	Db	6301	AGCAATACTACATACACAACAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAA	6360
	Qy	6361	ATAGTCGCGCTACCCCTCAGSAAATTTTTTTGTCTGTGTGTA CCTACGCTATCGTGTGTTG	6420
	Db	6361	ATAGTCGCGCTACCCCTCAGSAAATTTTTTTGTCTGTGTGTA CCTACGCTATCGTGTGTTG	6420
	Qy	6421	AATGGCTCTTCAGAACTATGTGCTCTCTCATCTTAGTGCCCCCTCATGRCACTAC	6480
	Db	6421	AATGGCTCTTCAGAACTATGTGCTCTCTCATCTTAGTGCCCCCTCATGRCACTAC	6480
	Qy	6481	ACTGAA CAAGATTTATACAGTTATGTCAATCTTAAGCCCGCCGAAACAAAAGAGTACCCA TT	6540
	Db	6481	ACTGAA CAAGATTTATACAGTTATGTCAATCTTAAGCCCGCCGAAACAAAAGAGTACCCA TT	6540
	Qy	6541	CTTCCCTTTGTTATAGAGCAGAGTGCTAGTGGCACTAGGTACTGGCAATGGCGGTATC	6600
	Db	6541	CTTCCCTTTGTTATAGAGCAGAGTGCTAGTGGCACTAGGTACTGGCAATGGCGGTATC	6600
	Qy	6601	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	6660
	Db	6601	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	6660
	Qy	6661	GTCGCGACTCCCTGGTCACTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT	6720
	Db	6661	GTCGCGACTCCCTGGTCACTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT	6720
	Qy	6721	CRAAATCGAAGACTTTAGACTVGYCTAAACCGCTGARAGAGGGGAAACCTGTTTATTTTTA	6780
	Db	6721	CRAAATCGAAGACTTTAGACTVGYCTAAACCGCTGARAGAGGGGAAACCTGTTTATTTTTA	6780
	Qy	6781	GGGGAAGATGCTGTTATTATGTTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAA	6840
	Db	6781	GGGGAAGATGCTGTTATTATGTTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAA	6840
	Qy	6841	ATTCSAGATCGAATA CAACGTAKAGCAGARGAGCTTCGAAAACACTGGACCTCGGGCCCTC	6900
	Db	6841	ATTCSAGATCGAATA CAACGTAKAGCAGARGAGCTTCGAAAACACTGGACCTCGGGCCCTC	6900
	Qy	6901	CTCAGCCCATPGGATGCCCTGGATTTCTCCCTTTTAGGACCTCTAGCAGCTATAATTTG	6960

	Dn	6901	CTCAGCCRATGGATGCCGTGATTCTTCTCCCCTTCCTTAGGACCTCTAGCAGCTATAAATTTGC	6986
	Qy	6961	CTACTCTCTTTTGGAACCTCGTAATCTTTTRACCCTCTCTTTAACTTTTGTCTCTCCAGAATC	7020
	Dn	6961	CTACTCTCTTTTGGAACCTCGTAATCTTTTRACCCTCTCTTTAACTTTTGTCTCTCCAGAATC	7020
	Qy	7021	GAACTGTAAAACTACAATTTGGAGCCCAAGATGCAGTCCAAGACTAAGATCTTACCGCAGA	7080
	Dn	7021	GAACTGTAAAACTACAATTTGGAGCCCAAGATGCAGTCCAAGACTAAGATCTTACCGCAGA	7080
	Qy	7081	CCCTTGACCGGCTGYTAGCCACATCTGATTTTAATGACATCAAAGAGCACCCCTCTCT	7140
	Dn	7081	CCCTTGACCGGCTGYTAGCCACATCTGATTTTAATGACATCAAAGAGCACCCCTCTCT	7140
	Qy	7141	GAGAAATCTCAGCTGCACAACCTCTACTACGCCCAATTTTCTGTGAGATGGGAGCTGAGA	7200
	Dn	7141	GAGAAATCTCAGCTGCACAACCTCTACTACGCCCAATTTTCTGTGAGATGGGAGCTGAGA	7200
	Qy	7201	GTSSTGCGGCAACCTCCCCAACAGCACCTTAGGTTTTCTGTTGAGATGGGAGCTGAGA	7260
	Dn	7201	GTSSTGCGGCAACCTCCCCAACAGCACCTTAGGTTTTCTGTTGAGATGGGAGCTGAGA	7260
	Qy	7261	GACAGACTAGCTGGATTTCTTAGCTGAYTAAGAATCCYTAAGCTTAGSGTGGAAAGTGT	7320
	Dn	7261	GACAGACTAGCTGGATTTCTTAGCTGAYTAAGAATCCYTAAGCTTAGSGTGGAAAGTGT	7320
	Qy	7321	ACCACATCCACCTTTAAAACACGCGGGCTGCACTTAGYTACACCTGACCAATTCAGAG	7380
	Dn	7321	ACCACATCCACCTTTAAAACACGCGGGCTGCACTTAGYTACACCTGACCAATTCAGAG	7380
	Qy	7381	CTCACTAAAATGCTAATTAGGCAAGACAGAGGTAAGAATAATAGCCAATCATYTTATTC	7440
	Dn	7381	CTCACTAAAATGCTAATTAGGCAAGACAGAGGTAAGAATAATAGCCAATCATYTTATTC	7440
	Qy	7441	MTGAGAGCAGCAGAGAGGACAAATGATCGGATATAAACCCCAAGTYTTGCGACCGGCAA	7500
	Dn	7441	MTGAGAGCAGCAGAGAGGACAAATGATCGGATATAAACCCCAAGTYTTGCGACCGGCAA	7500
	Qy	7501	CSCCAACCCCTTTGGGTCCTCCCTTGTATGGAGCTCTGTTTTTATGCTATTTTCAC	7560
	Dn	7501	CSCCAACCCCTTTGGGTCCTCCCTTGTATGGAGCTCTGTTTTTATGCTATTTTCAC	7560
	Qy	7561	TCTATTAAATCTTGCACTGCR	7582
	Dn	7561	TCTATTAAATCTTGCACTGCR	7582
	RESULT 2	BD267487	7582 bp DNA linear	PAT 17-JUL-2003
	LOCUS	DEFINITION	Endogenetic nucleic fragment associated with an autoimmune disease,	
	ACCESSION		marking method and reagent.	
	VERSION	BD267487		
	KEYWORDS	BD267487.1 GI:33077255		
	SOURCE	JP 2002534980-A/28.		
	ORGANISM	Homo sapiens (human)		
		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheraia; Euarchontoglires; Primates; Catarrhini;		
		Hominidae; Homo.		
	REFERENCE	1 (bases 1 to 7582)		
	AUTHORS	Baccala,G.P., Mallet,F. and Voisset,C.		
	TITLE	Endogenetic nucleic fragment associated with an autoimmune disease,		
	JOURNAL	marking method and reagent		
		Patent: JP 2002534980-A 28 22-OCT-2002;		
	COMMENT	BIO MERIEUX		
		OS Homo sapiens (human)		
		PN JP 2002534980-A/28		
		PD 22-OCT-2002		
		PF 21-JAN-2000 JP 2000594929		
		PR 21-JUN-1999 FR 99/00888		
		PI GLAUCIA PARANHOS BACCALA, FRANCOIS MALLET, CECILE VOISSET PC		
		C12N15/09.C07K14/15.C12O1/02.C12O1/68.G01N33/53.G01N33/53 DC		

G01N33/564,  
 PC G01N33/566 G01N33/569 G01N33/58//C12P21/02,C12N15/00 CC  
 Endogenous nucleic fragment associated with an autoimmune CC  
 disease,  
 CC marking method and reagent  
 FH Key Location/Qualifiers  
 FT source 1..7582  
 PT /organism='Homo sapiens (human)'.  
 FEATURES  
 source  
 1..7582  
 /organism='Homo sapiens'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:9606'  
 ORIGIN  
 Query Match 100.0%; Score 7582; DB 2; Length 7582;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAACAATCGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
 DB 1 CAACAATCGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
 QY 61 CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120  
 DB 61 CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCAACTGCA 120  
 QY 121 CTCCTTCGTGTCATGTTTCTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACACTGC 180  
 DB 121 CTCCTTCGTGTCATGTTTCTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACACTGC 180  
 QY 181 TGTTTGGCACACCGGANNACCTGCGCTGACTCCCATCCCTCTGATCTCGAGGGTGC 240  
 DB 181 TGTTTGGCACACCGGANNACCTGCGCTGACTCCCATCCCTCTGATCTCGAGGGTGC 240  
 QY 241 CGCTGTGCTTCATGATCCAGCGARGCGCCATTCGCGCTCCCAATTTGGGTAAAGGCTTGC 300  
 DB 241 CGCTGTGCTTCATGATCCAGCGARGCGCCATTCGCGCTCCCAATTTGGGTAAAGGCTTGC 300  
 QY 301 CATTTGNCCTGCACGGCTAAAGTGCCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360  
 DB 301 CATTTGNCCTGCACGGCTAAAGTGCCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360  
 QY 361 GGGTTCATGGTCTCTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTTACCACA 420  
 DB 361 GGGTTCATGGTCTCTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTTACCACA 420  
 QY 421 TGGCCCAAGATTCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTCAAGAAATAC 480  
 DB 421 TGGCCCAAGATTCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTCAAGAAATAC 480  
 QY 481 GARGCTTGCACCATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC 540  
 DB 481 GARGCTTGCACCATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC 540  
 QY 541 TTGGGAGCTCTGTGACGAAGAACCCCGGTACATTTTGGGACACMSRACGGACATCC 600  
 DB 541 TTGGGAGCTCTGTGACGAAGAACCCCGGTACATTTTGGGACACMSRACGGACATCC 600  
 QY 601 MAAGTGTATGGGAAAGCTTCCCGCAAGACAAAAACGCGCTTAAGACGTATTCTGGARAAT 660  
 DB 601 MAAGTGTATGGGAAAGCTTCCCGCAAGACAAAAACGCGCTTAAGACGTATTCTGGARAAT 660  
 QY 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTCTGCAAGTGC 720  
 DB 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTCTGCAAGTGC 720  
 QY 721 GCCTGGCACTCTGAGGGAAGTATAAATTTAATACCATCTTACAGCTAGACTCTTTTG 780  
 DB 721 GCCTGGCACTCTGAGGGAAGTATAAATTTAATACCATCTTACAGCTAGACTCTTTTG 780  
 QY 781 TAGAAAAAGGCAATGGAGTGAAGTGCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA 840  
 DB 781 TAGAAAAAGGCAATGGAGTGAAGTGCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA 840

DB 781 TAGAAAAAGGCAATGGAGTGAAGTGCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA 840  
 QY 841 CTCACAAATTATGTAATAAAGTGTGATTTATGCCCCCTACAGGAAGCCTTACAGTCTACCTCC 900  
 DB 841 CTCACAAATTATGTAATAAAGTGTGATTTATGCCCCCTACAGGAAGCCTTACAGTCTACCTCC 900  
 QY 901 CTATCCCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCAAATGG 960  
 DB 901 CTATCCCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCAAATGG 960  
 QY 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACAAAGAGTGCCCAATATCCCAAT 1020  
 DB 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACAAAGAGTGCCCAATATCCCAAT 1020  
 QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGCCCCCAGCAGAGTGCATGTGCT 1080  
 DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGCCCCCAGCAGAGTGCATGTGCT 1080  
 QY 1081 TTTTCTTCCCGACTTTAAAGCAATTAAGCAAGACTTTAGGTAAATCTCAGATAAATGCT 1140  
 DB 1081 TTTTCTTCCCGACTTTAAAGCAATTAAGCAAGACTTTAGGTAAATCTCAGATAAATGCT 1140  
 QY 1141 GATGGCTATATTTGTTTATCAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAT 1200  
 DB 1141 GATGGCTATATTTGTTTATCAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAT 1200  
 QY 1201 ATATATGTCATGCTTAATCAGACACTAAACCCCAATTAAGGAAGTGCACATAACTGC 1260  
 DB 1201 ATATATGTCATGCTTAATCAGACACTAAACCCCAATTAAGGAAGTGCACATAACTGC 1260  
 QY 1261 AGCTGTAGGTTTGGCGATCTCTGATCTCAGTCAGTCAATGATGATGATGATGATGATGAT 1320  
 DB 1261 AGCTGTAGGTTTGGCGATCTCTGATCTCAGTCAGTCAATGATGATGATGATGATGATGAT 1320  
 QY 1321 GAAGAAAGANAATGATTTCCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1380  
 DB 1321 GAAGAAAGANAATGATTTCCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1380  
 QY 1381 GGGACACAGAAATCAGTAAACATGGGAGATTGGTGTGAGACATTTGCTTAATCTGTGTGC 1440  
 DB 1381 GGGACACAGAAATCAGTAAACATGGGAGATTGGTGTGAGACATTTGCTTAATCTGTGTGC 1440  
 QY 1441 TASAAGGACTAAGGAATACTASGAAGAAATCTAYGAATTAATCAATGATGTCACCAATA 1500  
 DB 1441 TASAAGGACTAAGGAATACTASGAAGAAATCTAYGAATTAATCAATGATGTCACCAATA 1500  
 QY 1501 ACACAGGGAAGGAAGAAATCTCTACTGCTTTCTGAGAGACTTAAGGAGGCAATTGAG 1560  
 DB 1501 ACACAGGGAAGGAAGAAATCTCTACTGCTTTCTGAGAGACTTAAGGAGGCAATTGAG 1560  
 QY 1561 GAAGGTGCTCTCTGTGACCTGACTCTTTGAAAGGCCAATAATCTTTAAAGCGTAAGTT 1620  
 DB 1561 GAAGGTGCTCTCTGTGACCTGACTCTTTGAAAGGCCAATAATCTTTAAAGCGTAAGTT 1620  
 QY 1621 TATCACTCAGTCAGTGCAGACATTAAGAAATCTTCAAAAGTCTGCGTAGGCCCGGAG 1680  
 DB 1621 TATCACTCAGTCAGTGCAGACATTAAGAAATCTTCAAAAGTCTGCGTAGGCCCGGAG 1680  
 QY 1681 CAAAACCTTAGAAACCTTATGAACTGCGTCTGCGTCTTTTATTAATAGAGATCAGGAG 1740  
 DB 1681 CAAAACCTTAGAAACCTTATGAACTGCGTCTGCGTCTTTTATTAATAGAGATCAGGAG 1740  
 QY 1741 GAGCAGCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTT 1800  
 DB 1741 GAGCAGCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTT 1800  
 QY 1801 CAGCAGTGCATTTGGAGGCTCTGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAG 1860  
 DB 1801 CAGCAGTGCATTTGGAGGCTCTGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAG 1860  
 QY 1861 TAGGGCTTGTCTTCCAGTGGGTCTACAGGACACTTTTAAAAAAGATTTCTCAAGTAGAAG 1920  
 DB 1861 TAGGGCTTGTCTTCCAGTGGGTCTACAGGACACTTTTAAAAAAGATTTCTCAAGTAGAAG 1920

QY 1921 TAAGCGCCCTTGGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
DB 1921 TAAGCGCCCTTGGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
QY 1981 CAGGGACAAGAGTCTTTTGGAGTCAGAAGCACTAACAGATGATCCAGCAGCAGACTG 2040  
DB 1981 CAGGGACAAGAGTCTTTTGGAGTCAGAAGCACTAACAGATGATCCAGCAGCAGACTG 2040  
QY 2041 AGGGTGGTGGGCAAGCGCCATCCCATGCCATCACTCCAGAGCCCTGGGTATGCTT 2100  
DB 2041 AGGGTGGTGGGCAAGCGCCATCCCATGCCATCACTCCAGAGCCCTGGGTATGCTT 2100  
QY 2101 GACCATTTAGGGCCAGGAAGTGTCTCTGGACACTGTGGCGGTCTTCTAGTCTTACT 2160  
DB 2101 GACCATTTAGGGCCAGGAAGTGTCTCTGGACACTGTGGCGGTCTTCTAGTCTTACT 2160  
QY 2161 CTTCTGTCCCGGCAAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGAGC 2220  
DB 2161 CTTCTGTCCCGGCAAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGAGC 2220  
QY 2221 GGCACTCATAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280  
DB 2221 GGCACTCATAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280  
QY 2281 TCACATGCTTTTCTAATTTATGCTTGAAGCCCCCACTACCTTGTGTGGGAGACATTTCTA 2340  
DB 2281 TCACATGCTTTTCTAATTTATGCTTGAAGCCCCCACTACCTTGTGTGGGAGACATTTCTA 2340  
QY 2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTCTGTNCC 2400  
DB 2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTCTGTNCC 2400  
QY 2401 CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGCAATATGGACGAGC 2460  
DB 2401 CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGCAATATGGACGAGC 2460  
QY 2461 CAAGAATGCCGTCTCTGTTCAAGTTAACTAAGGATTCACCTTCTTCCCTACCAAA 2520  
DB 2461 CAAGAATGCCGTCTCTGTTCAAGTTAACTAAGGATTCACCTTCTTCCCTACCAAA 2520  
QY 2521 GGCAATACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAGATTGTTAAGCACTTA 2580  
DB 2521 GGCAATACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAGATTGTTAAGCACTTA 2580  
QY 2581 AAGCCCAAGGCTTAGTAAACCATGCACTCCCTGCAGTAATTCGGTAGTGGATTGAG 2640  
DB 2581 AAGCCCAAGGCTTAGTAAACCATGCACTCCCTGCAGTAATTCGGTAGTGGATTGAG 2640  
QY 2641 GAGGCAAGAAACCCAGTGGAAGTGAGGAGGTTAGTGCAAGATCTCAGGATTAATCAATG 2700  
DB 2641 GAGGCAAGAAACCCAGTGGAAGTGAGGAGGTTAGTGCAAGATCTCAGGATTAATCAATG 2700  
QY 2701 AGGCCGTGTCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCTCAATAC 2760  
DB 2701 AGGCCGTGTCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCTCAATAC 2760  
QY 2761 CAGAGGAAGCAGAGTGTTTACASTCTCGACCTTMAAGATGCTTCTTCTGATCCCTG 2820  
DB 2761 CAGAGGAAGCAGAGTGTTTACASTCTCGACCTTMAAGATGCTTCTTCTGATCCCTG 2820  
QY 2821 TACATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATACTTCAAAACCCCACTCAAC 2880  
DB 2821 TACATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATACTTCAAAACCCCACTCAAC 2880  
QY 2881 TCACCTGGACTRTTTTACCCCAAGGTTCAAGGATAGYCCCATCTATTTTGGCAGGAT 2940  
DB 2881 TCACCTGGACTRTTTTACCCCAAGGTTCAAGGATAGYCCCATCTATTTTGGCAGGAT 2940  
QY 2941 TAGCCCAAGACTTGAGYCAATYMTACATCTGACACTCTTCTCTCRGTAKGTGGATG 3000  
DB 2941 TAGCCCAAGACTTGAGYCAATYMTACATCTGACACTCTTCTCTCRGTAKGTGGATG 3000

QY 3001 ATTTACTTTTRGCGCYGCTTTTCAGAAACCTTTGTGCCATCAAGCCACCCAGCRCTCTTMA 3060  
DB 3001 ATTTACTTTTRGCGCYGCTTTTCAGAAACCTTTGTGCCATCAAGCCACCCAGCRCTCTTMA 3060  
QY 3061 ATTTCTCTCGYACCTGTGGCTACAWGTTTCCAAACSAARAGCTCARCTCTGTCTCACAGC 3120  
DB 3061 ATTTCTCTCGYACCTGTGGCTACAWGTTTCCAAACSAARAGCTCARCTCTGTCTCACAGC 3120  
QY 3121 AGGTTAAATACCTTAGGRCCTAARATTTATCCAAAGGCCACCAAGGCCCTCAGTGAGGAAYRYA 3180  
DB 3121 AGGTTAAATACCTTAGGRCCTAARATTTATCCAAAGGCCACCAAGGCCCTCAGTGAGGAAYRYA 3180  
QY 3181 TCCAGCCTATATCTGGCTTATCTCATCYCAAAACCCCTAAGCAACTAAGGRRTTCTTTCG 3240  
DB 3181 TCCAGCCTATATCTGGCTTATCTCATCYCAAAACCCCTAAGCAACTAAGGRRTTCTTTCG 3240  
QY 3241 GCRTAAYAGGYTTCTGCCGAATGATTTCCCGAGTGGGCAATAGCCAGGYCATTTA 3300  
DB 3241 GCRTAAYAGGYTTCTGCCGAATGATTTCCCGAGTGGGCAATAGCCAGGYCATTTA 3300  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCCAATACCCATTTTARTAGATGGAYAMCTGAA 3360  
DB 3301 WATACASTAATTAAGGAACTCAGAAAGCCAATACCCATTTTARTAGATGGAYAMCTGAA 3360  
QY 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCCAAGTGTTAAGYT 3420  
DB 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCCAAGTGTTAAGYT 3420  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTATATVTCACAGAAACCAAGAAAGAGTGTCTRGA 3480  
DB 3421 TGCCAAACRGGGCAAGACTTTTSTTATATVTCACAGAAACCAAGAAAGAGTGTCTRGA 3480  
QY 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTTCACACCTTTCGCRVACCTGASTAAGGAAAYT 3540  
DB 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTTCACACCTTTCGCRVACCTGASTAAGGAAAYT 3540  
QY 3541 GATGTAGTGGCAAGGTTGRCYTCATTTGTTTAYGGTGTGGTGGCAGTAGCAGTYKTA 3600  
DB 3541 GATGTAGTGGCAAGGTTGRCYTCATTTGTTTAYGGTGTGGTGGCAGTAGCAGTYKTA 3600  
QY 3601 GTATCTGAAGCAGTTTAAATATACAGGAGAGATCTTACTGTGGACATCTCATGAK 3660  
DB 3601 GTATCTGAAGCAGTTTAAATATACAGGAGAGATCTTACTGTGGACATCTCATGAK 3660  
QY 3661 GTGAAYRGCACTACTCTCTCTAAAGAGACTTGTGGCTGTGACAAACCTGTTTACTTAAA 3720  
DB 3661 GTGAAYRGCACTACTCTCTCTAAAGAGACTTGTGGCTGTGACAAACCTGTTTACTTAAA 3720  
QY 3721 TRTCAGGCTCTATTAATTCTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
DB 3721 TRTCAGGCTCTATTAATTCTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
QY 3781 CCAGYCNCACTTCTTCCAGCACTAAGAAAGATARAAYATACTGTCAACAAATTAAT 3840  
DB 3781 CCAGYCNCACTTCTTCCAGCACTAAGAAAGATARAAYATACTGTCAACAAATTAAT 3840  
QY 3841 TCTCAAACTATGCCACTCGAGGGGACCTTYPAGARGTTCCYTTGACCTGATCCYACCTT 3900  
DB 3841 TCTCAAACTATGCCACTCGAGGGGACCTTYPAGARGTTCCYTTGACCTGATCCYACCTT 3900  
QY 3901 CAACTTGTATCTGTATGGAAGTTCTTTGTAGAAAAGGACTTCCGAAAGYGGGATGTC 3960  
DB 3901 CAACTTGTATCTGTATGGAAGTTCTTTGTAGAAAAGGACTTCCGAAAGYGGGATGTC 3960  
QY 3961 AGTGTCTAGTATTAATGAATATTTGAAGTAATCCCTCTACTCAGGAACTAGTGTCTYA 4020  
DB 3961 AGTGTCTAGTATTAATGAATATTTGAAGTAATCCCTCTACTCAGGAACTAGTGTCTYA 4020  
QY 4021 GCTCCGCAACTTAATAGCCYCACTKGGCCTAGAAATAGGAGAGGAAAGGGYAAA 4080  
DB 4021 GCTCCGCAACTTAATAGCCYCACTKGGCCTAGAAATAGGAGAGGAAAGGGYAAA 4080  
QY 4081 TATATATACAGACTCTRTARTATGCTYACCTAGTCTNTCCATGCCCATGMRCAATATGSAR 4140

|||||  
4081 TATATATACAGACTCTIRARTATGCTVACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
|||||  
4141 AGAAGGGAATTCCTAACTTCCTGAGRGAAACACTATCAMACATCAGGAAGCCATTAGGAR 4200  
|||||  
4141 AGAAGGGAATTCCTAACTTCCTGAGRGAAACACTATCAMACATCAGGAAGCCATTAGGAR 4200  
|||||  
4201 ATTATTATYGGCGWTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCGGGGTCAATCA 4260  
|||||  
4201 ATTATTATYGGCGWTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCGGGGTCAATCA 4260  
|||||  
4261 NAAAGGAAAGRAAGGGAATASAGRGAAATGCGCAAGCAKATATTGAAGCMAAAGAGC 4320  
|||||  
4261 NAAAGGAAAGRAAGGGAATASAGRGAAATGCGCAAGCAKATATTGAAGCMAAAGAGC 4320  
|||||  
4321 TCGAAGGCAAGGACCTTCCTAATAGAAATGCTTATTAATCTTCCCTTAGTATAGGGTAATCC 4380  
|||||  
4321 TCGAAGGCAAGGACCTTCCTAATAGAAATGCTTATTAATCTTCCCTTAGTATAGGGTAATCC 4380  
|||||  
4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCACGAGG 4440  
|||||  
4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCACGAGG 4440  
|||||  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTCCTGCGCAAC 4500  
|||||  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTCCTGCGCAAC 4500  
|||||  
4501 TATCCAATGGAAATTACTTAAACCTTCATCAAACTTTCTCATTAGGCATCGATAGCAC 4560  
|||||  
4501 TATCCAATGGAAATTACTTAAACCTTCATCAAACTTTCTCATTAGGCATCGATAGCAC 4560  
|||||  
4561 CCATCARATGGCCAAATCAATTAATTTACTGGACAGGCGCTTTTCAAAACTATCAAGGARAT 4620  
|||||  
4561 CCATCARATGGCCAAATCAATTAATTTACTGGACAGGCGCTTTTCAAAACTATCAAGGARAT 4620  
|||||  
4621 AKTCAGGCGCTGTGAATCTGCCARAAATAATCCCTGCGCTYATGCGCAAGCTCCTTTC 4680  
|||||  
4621 AKTCAGGCGCTGTGAATCTGCCARAAATAATCCCTGCGCTYATGCGCAAGCTCCTTTC 4680  
|||||  
4681 AGGARAACAAARAACAGGCCATTACCCCTGRARAARACTGGCAACTGATTTTACCACAAG 4740  
|||||  
4681 AGGARAACAAARAACAGGCCATTACCCCTGRARAARACTGGCAACTGATTTTACCACAAG 4740  
|||||  
4741 CCCAAACCTCAGGGATTTTCAAGTACTTACTAGTCTGGGTARATACCTTACGGGTTGGGCA 4800  
|||||  
4741 CCCAAACCTCAGGGATTTTCAAGTACTTACTAGTCTGGGTARATACCTTACGGGTTGGGCA 4800  
|||||  
4801 RAGGCGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTATGAA 4860  
|||||  
4801 RAGGCGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTATGAA 4860  
|||||  
4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
|||||  
4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
|||||  
4921 GCCACAGTAAACCGAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGGCC 4980  
|||||  
4921 GCCACAGTAAACCGAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGGCC 4980  
|||||  
4981 TGAAGGCCACAGTCTCAGGGAAGTTCGAGAAATAAGAACTCAAAAGGACATCTA 5040  
|||||  
4981 TGAAGGCCACAGTCTCAGGGAAGTTCGAGAAATAAGAACTCAAAAGGACATCTA 5040  
|||||  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGYTCTGTGTGCTATAGCCCTTAAAA 5100  
|||||  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGYTCTGTGTGCTATAGCCCTTAAAA 5100  
|||||  
5101 AGAATCTGAACCTTTCCCAAAAAGCAGGACTTAGCCCATCGAATGCTGTATGGGAAG 5160  
|||||  
5101 AGAATCTGAACCTTTCCCAAAAAGCAGGACTTAGCCCATCGAATGCTGTATGGGAAG 5160  
|||||  
5161 CCTTTCATACCAATGACCTTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGACAGATCA 5220  
|||||

Db 5161 CCTTTATAACCAATGACCTTGTGCTTGACCCAAAGACAGCCAACTTAGTTGACAGATCA 5220  
Qy 5221 CCTCTTAGCCAAATATCAACAAGTTCTTAAACAATTAAGGAACCTTATCCCTCAGAG 5280  
|||||  
Db 5221 CCTCTTAGCCAAATATCAACAAGTTCTTAAACAATTAAGGAACCTTATCCCTCAGAG 5280  
Qy 5281 AGGGAAGAAACTATTTCCACCCWMTGACATGCTATTTAGTCAAGTCCCTTCTCTTAAT 5340  
|||||  
Db 5281 AGGGAAGAAACTATTTCCACCCWMTGACATGCTATTTAGTCAAGTCCCTTCTCTTAAT 5340  
Qy 5341 CCCCATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATYATACCCCACTG 5400  
|||||  
Db 5341 CCCCATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATYATACCCCACTG 5400  
Qy 5401 CGGTTAAGTGGCTGGAGTGGAGTCTTGGAATACATCACACTTGAGTCAAACTCTGGATAC 5460  
|||||  
Db 5401 CGGTTAAGTGGCTGGAGTGGAGTCTTGGAATACATCACACTTGAGTCAAACTCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAAACCTCTAGAGG 5520  
|||||  
Db 5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAAACCTCTAGAGG 5520  
Qy 5521 ATTTGGCGCTCTCTTCAACAACAACAGGAGGAAGTAACATAAATCAATAATCCGCC 5580  
|||||  
Db 5521 ATTTGGCGCTCTCTTCAACAACAACAGGAGGAAGTAACATAAATCAATAATCCGCC 5580  
Qy 5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTCACTCTCACT 5640  
|||||  
Db 5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
|||||  
Db 5641 GCACCCCTCCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGACAGCTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCC 5760  
|||||  
Db 5701 ATGACAGCTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCC 5760  
Qy 5761 ACCTTCACTGCCCCACCCCATATGCCCCGGAACCTGCTATCACTGCCCCCTTTTGCATG 5820  
|||||  
Db 5761 ACCTTCACTGCCCCACCCCATATGCCCCGGAACCTGCTATCACTGCCCCCTTTTGCATG 5820  
Qy 5821 CATGCAAACTACTCATTTATGGACAGGAAATAATGATTAATCTAGTTGCTCTGGAGACTT 5880  
|||||  
Db 5821 CATGCAAACTACTCATTTATGGACAGGAAATAATGATTAATCTAGTTGCTCTGGAGACTT 5880  
Qy 5881 GGAGTCACTGCTGTGTGGACTTACTTCAACCACAACTGGTATGCTGTATGGGGGTGGAGTT 5940  
|||||  
Db 5881 GGAGTCACTGCTGTGTGGACTTACTTCAACCACAACTGGTATGCTGTATGGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGGAAATAATGTAAGAAGTAATCTCCCAACTCACCGGGGTACT 6000  
|||||  
Db 5941 CAAGATCAGGCAAGGAAATAATGTAAGAAGTAATCTCCCAACTCACCGGGGTACT 6000  
Qy 6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACC 6060  
|||||  
Db 6001 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACC 6060  
Qy 6061 CATACTCGCCTGGTAAGCTATTTTAATACCACTCTCTGCGGCTCCATGAGGTCTCGGCC 6120  
|||||  
Db 6061 CATACTCGCCTGGTAAGCTATTTTAATACCACTCTCTGCGGCTCCATGAGGTCTCGGCC 6120  
Qy 6121 CAAAACCTCTACTGTTGGATATGCTCTCCCTGAACTTCARGGCATATGTTTCAATC 6180  
|||||  
Db 6121 CAAAACCTCTACTGTTGGATATGCTCTCCCTGAACTTCARGGCATATGTTTCAATC 6180  
Qy 6181 CCGTGTACCTGAACCAATGGAACAACTTCAGCAAGAAATAAACCACTTCCGTTTATGTA 6240  
|||||  
Db 6181 CCGTGTACCTGAACCAATGGAACAACTTCAGCAAGAAATAAACCACTTCCGTTTATGTA 6240  
Qy 6241 GGACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCCTGTTGTAATAATTT 6300  
|||||  
Db 6241 GGACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCCTGTTGTAATAATTT 6300



QY 6301 AGCAATACTACATACACAACTCCCAATGCATCAGGTGGTAACTCTCTCCACACAA 6360  
 Db 6301 AGCAATACTACATACACAACTCCCAATGCATCAGGTGGTAACTCTCTCCACACAA 6360  
 QY 6361 ATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTG 6420  
 Db 6361 ATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTG 6420  
 QY 6421 AATGGCTCTTCAGAACTATGTGCTCTCAATCTTATAGTGCCTCCATGRCATCTAC 6480  
 Db 6421 AATGGCTCTTCAGAACTATGTGCTCTCAATCTTATAGTGCCTCCATGRCATCTAC 6480  
 QY 6481 ACTGAACAAGATTTATACAGTTATGTATCATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
 Db 6481 ACTGAACAAGATTTATACAGTTATGTATCATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
 QY 6541 CTTCTCTTTTGTATAGGAGCAGGAGTCTAGGTGCATCTAGTACTGGCATTTGGCGGTATC 6600  
 Db 6541 CTTCTCTTTTGTATAGGAGCAGGAGTCTAGGTGCATCTAGTACTGGCATTTGGCGGTATC 6600  
 QY 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
 Db 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
 QY 6661 GTCCGCGACTCCCTGCTCAGCTTGCAGATCACTTAACTCCCTAGCAGTAGTCCCTT 6720  
 Db 6661 GTCCGCGACTCCCTGCTCAGCTTGCAGATCACTTAACTCCCTAGCAGTAGTCCCTT 6720  
 QY 6721 CRAAATCGAAGACTTTAGACTTGTAACTGCTGAGAGGGGAACTGTTTATTTTAA 6780  
 Db 6721 CRAAATCGAAGACTTTAGACTTGTAACTGCTGAGAGGGGAACTGTTTATTTTAA 6780  
 QY 6781 GGGGAAGAATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
 Db 6781 GGGGAAGAATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
 QY 6841 ATTCSAGATCGAATACAACTGATKAGCAGAGAGCTTCGAAACACTCGACCTGGGCGCTC 6900  
 Db 6841 ATTCSAGATCGAATACAACTGATKAGCAGAGAGCTTCGAAACACTCGACCTGGGCGCTC 6900  
 QY 6901 CTCAGCCGATGATGCTGCTGCTTCTTACCTCTTGTAACTTGTCTCTCCAGAAATC 7020  
 Db 6901 CTCAGCCGATGATGCTGCTGCTTCTTACCTCTTGTAACTTGTCTCTCCAGAAATC 7020  
 QY 7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGA 7080  
 Db 7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGA 7080  
 QY 7081 CCCCTGGACGGCTGTATAGCCCAAGATCTGATGTTAATGACATCAAGGCAACCTCTCT 7140  
 Db 7081 CCCCTGGACGGCTGTATAGCCCAAGATCTGATGTTAATGACATCAAGGCAACCTCTCTCT 7140  
 QY 7141 GAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGAGTATAGAGC 7200  
 Db 7141 GAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGAGTATAGAGC 7200  
 QY 7201 GGTSGTGGCCAACTCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGA 7260  
 Db 7201 GGTSGTGGCCAACTCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGA 7260  
 QY 7261 GACAGACTAGCTGGATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAGGTG 7320  
 Db 7261 GACAGACTAGCTGGATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAGGTG 7320  
 QY 7321 ACCACATCCACCTTTTAAACACGGGGCTGCACTTAGTTCACACCTGACCAATCAGAGAG 7380  
 Db 7321 ACCACATCCACCTTTTAAACACGGGGCTGCACTTAGTTCACACCTGACCAATCAGAGAG 7380

QY 7381 CTCACATAAATCTAATAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTGC 7440  
 Db 7381 CTCACATAAATCTAATAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTGC 7440  
 QY 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCACCAAGTTCGAGCCGCGCAA 7500  
 Db 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCACCAAGTTCGAGCCGCGCAA 7500  
 QY 7501 CGGCAACCCCTTTGGGTCCCTTCTGATGAGGAGCTCTGTTTTCATGCTATTTCAC 7560  
 Db 7501 CGGCAACCCCTTTGGGTCCCTTCTGATGAGGAGCTCTGTTTTCATGCTATTTCAC 7560  
 QY 7561 TCTATTAAATCTTGCARCTGCR 7582  
 Db 7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 3  
 AX000966  
 LOCUS Sequence 11 from Patent WO9902696.  
 DEFINITION AX000966  
 ACCESSION AX000966  
 VERSION AX000966.1 GI:7241208  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unclassified  
 unclassified sequences.  
 REFERENCE 1 (bases 1 to 7582)  
 AUTHORS Beseme, F. and Blond, J.  
 TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE  
 DISEASES OR WITH PREGNANCY DISORDERS  
 JOURNAL Patent: WO 9902696-A 11 21-JAN-1999;  
 BIO MERIEUX (FR); BESEME FREDERIC (FR)  
 FEATURES  
 Location/Qualifiers  
 1..7582  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

ORIGIN  
 Query Match 100.0%; Score 7582; DB 2; Length 7582;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAATCGGATATAAACCACCAAGGCAATTCGAGCTGGCAACAGCAGCCCTCTTGGGTGC 60  
 Db 1 CAACAATCGGATATAAACCACCAAGGCAATTCGAGCTGGCAACAGCAGCCCTCTTGGGTGC 60  
 QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACCTCTATTAATCTTGCACACTGCA 120  
 Db 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACCTCTATTAATCTTGCACACTGCA 120  
 QY 121 CTCCTTCTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180  
 Db 121 CTCCTTCTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180  
 QY 181 TGTGTTGCCACCAACCCGCAACACTGCGCTGACTCCCATCCCTGATCTCTGAGGCTGTC 240  
 Db 181 TGTGTTGCCACCAACCCGCAACACTGCGCTGACTCCCATCCCTGATCTCTGAGGCTGTC 240  
 QY 241 CGCTGCTCTCTGATCCAGCGARGCGCCCATTTGCCGCTCCCAATTTGGGCTAAAGGCTTGC 300  
 Db 241 CGCTGCTCTCTGATCCAGCGARGCGCCCATTTGCCGCTCCCAATTTGGGCTAAAGGCTTGC 300  
 QY 301 CATTTGTCCTGACCGGCTAAGTGCCTGGGTTGTTCTTAATGAGCTGAACACTANTCACT 360  
 Db 301 CATTTGTCCTGACCGGCTAAGTGCCTGGGTTGTTCTTAATGAGCTGAACACTANTCACT 360  
 QY 361 GGGTTCCTCATGTTCTTCTGTGACCCACGCTCTTAATAAACAATTAACACTTACCACA 420  
 Db 361 GGGTTCCTCATGTTCTTCTGTGACCCACGCTCTTAATAAACAATTAACACTTACCACA 420  
 QY 421 TGGCCCAAGATTCCATTCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480

|||||  
421 TGGCCCAAGATTCATTCTTGGAAATCCGTGAGGSCAAGAACTCCAGGTCAAGAAATAC 480  
QY  
481 GARGCTTGCCACCAATCTTTGGAAAGCGCCTGTACCTCTTTGGAAAGTGGTTCAACCAACATC 540  
Db  
481 GARGCTTGCCACCAATCTTTGGAAAGCGCCTGTACCTCTTTGGAAAGTGGTTCAACCAACATC 540  
QY  
541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGACCAACNSRACGGACATCC 600  
Db  
541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGACCAACNSRACGGACATCC 600  
QY  
601 MAAGTGATGGGAAAGCTTCCCGCAAGACAGCAAAAACGCCCTTAAGACGTAATCTGGAAT 660  
Db  
601 MAAGTGATGGGAAAGCTTCCCGCAAGACAGCAAAAACGCCCTTAAGACGTAATCTGGAAT 660  
QY  
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATCTTCTGCAAGTGC 720  
Db  
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATCTTCTGCAAGTGC 720  
QY  
721 GCCTGGCACTCTGAGGGAAGTATAAATTTATAACCACTTTACAGCTAGACATCTTTTG 780  
Db  
721 GCCTGGCACTCTGAGGGAAGTATAAATTTATAACCACTTTACAGCTAGACATCTTTTG 780  
QY  
781 TAGAAAAGGCAATGGAGTGAAGTGCCATAAGTACAAAACCTTTCTTTTCAATTAAGAGACAA 840  
Db  
781 TAGAAAAGGCAATGGAGTGAAGTGCCATAAGTACAAAACCTTTCTTTTCAATTAAGAGACAA 840  
QY  
841 CTCACAATTTATGATAAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCC 900  
Db  
841 CTCACAATTTATGATAAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCC 900  
QY  
901 CTATCCAGCATCCCGACTCCTTCCCGAMYTAATAAGGACCCCTTCAACCCAAATGG 960  
Db  
901 CTATCCAGCATCCCGACTCCTTCCCGAMYTAATAAGGACCCCTTCAACCCAAATGG 960  
QY  
961 TCCAAAAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT 1020  
Db  
961 TCCAAAAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT 1020  
QY  
1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTTGGCCCAAGCAGAGTGCATGTGCT 1080  
Db  
1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTTGGCCCAAGCAGAGTGCATGTGCT 1080  
QY  
1081 TTTTCTCCAGACTTAAGGCAATATAAAGACAGCTTAGGTAAATTTCTCAGATTAATCT 1140  
Db  
1081 TTTTCTCCAGACTTAAGGCAATATAAAGACAGCTTAGGTAAATTTCTCAGATTAATCT 1140  
QY  
1141 GATGGCTATATTGRTGTTTTCAAGGTTTAGGCAATTTCTTTGATCTGACATGGAGAT 1200  
Db  
1141 GATGGCTATATTGRTGTTTTCAAGGTTTAGGCAATTTCTTTGATCTGACATGGAGAT 1200  
QY  
1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCCACCATACTGC 1260  
Db  
1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCCACCATACTGC 1260  
QY  
1261 AGCTGAGGRTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGATGATGATGACAAACA 1320  
Db  
1261 AGCTGAGGRTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGATGATGATGACAAACA 1320  
QY  
1321 GAAGGAAGAAATGATTTCCCAAGCCAGCAGCAGTTCACAGTCTASACCTCATTTG 1380  
Db  
1321 GAAGGAAGAAATGATTTCCCAAGCCAGCAGCAGTTCACAGTCTASACCTCATTTG 1380  
QY  
1381 GGGACACAGAAATCAGTAACATGGGAGATTTGGTGTGACAGACATTTGCTAACTTTGTGTC 1440  
Db  
1381 GGGACACAGAAATCAGTAACATGGGAGATTTGGTGTGACAGACATTTGCTAACTTTGTGTC 1440  
QY  
1441 TASAAGGACTAAGGAAAACTASGAAGAAATCTAYGAATTAATCTCAATGATGTCACCATTA 1500  
Db  
1441 TASAAGGACTAAGGAAAACTASGAAGAAATCTAYGAATTAATCTCAATGATGTCACCATTA 1500  
QY  
1501 ACACAGGGGAAGGGAAGAAATCTTACTGCTTTCTGGAGAGACTAAGGGAGGCAATTTGAG 1560

Db  
1501 ACACAGGGGAAGGGAAGAAATCTTACTGCTTTCTGGAGAGACTAAGGGAGGCAATTTGAG 1560  
QY  
1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620  
Db  
1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620  
QY  
1621 TATCACTCAGTCAGTCGAGACATTTAGAAAAAACTTCAAAAGTCTGCGTAGGCCCGGAG 1680  
Db  
1621 TATCACTCAGTCAGTCGAGACATTTAGAAAAAACTTCAAAAGTCTGCGTAGGCCCGGAG 1680  
QY  
1681 CAAAATTTAGAAACCTATTGAACCTTGGCAACCTTGGGTTTTTATAATAGAGATCAGAG 1740  
Db  
1681 CAAAATTTAGAAACCTATTGAACCTTGGCAACCTTGGGTTTTTATAATAGAGATCAGAG 1740  
QY  
1741 GAGCAGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTT 1800  
Db  
1741 GAGCAGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTT 1800  
QY  
1801 CAGGCAAGTGACCTTTGGAGGCTCTGGAAAAAGGAAAGCTGGGCAAAATTTGAATGCCTAA 1860  
Db  
1801 CAGGCAAGTGACCTTTGGAGGCTCTGGAAAAAGGAAAGCTGGGCAAAATTTGAATGCCTAA 1860  
QY  
1861 TAGGCTTTGCTTCCAGTGCGGTCTACAGGACACTTTAAAAAAGATTTGTTCAAGTAGAAG 1920  
Db  
1861 TAGGCTTTGCTTCCAGTGCGGTCTACAGGACACTTTAAAAAAGATTTGTTCAAGTAGAAG 1920  
QY  
1921 TAAGCGGCCCTTGTCTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
Db  
1921 TAAGCGGCCCTTGTCTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
QY  
1981 CAGGGCAAAAGGCTTTTGTAGTCAGAGCCACTAAACCAGATGATCCAGCAGCAGACTG 2040  
Db  
1981 CAGGGCAAAAGGCTTTTGTAGTCAGAGCCACTAAACCAGATGATCCAGCAGCAGACTG 2040  
QY  
2041 AGGCTGCTGGGCAAGCGCCATCCATGCCATCACTGAGAGCCCTTGGGTATGCTT 2100  
Db  
2041 AGGCTGCTGGGCAAGCGCCATCCATGCCATCACTGAGAGCCCTTGGGTATGCTT 2100  
QY  
2101 GACCATTTAGGGCCAGGAAGTTGTCTCTGGAACACTGGTGGGTCTTCTTAGTCTTACT 2160  
Db  
2101 GACCATTTAGGGCCAGGAAGTTGTCTCTGGAACACTGGTGGGTCTTCTTAGTCTTACT 2160  
QY  
2161 CTTCTGCTCCGCAACACTGTCTCCAGATCTGTCTACTTCTGAGGGGTCCNTAAGAGC 2220  
Db  
2161 CTTCTGCTCCGCAACACTGTCTCCAGATCTGTCTACTTCTGAGGGGTCCNTAAGAGC 2220  
QY  
2221 GGCAGTCACTAGATACTTTTCCAGCCACTAAAGTTATGAACCTGGGAGCTTTATTTCTT 2280  
Db  
2221 GGCAGTCACTAGATACTTTTCCAGCCACTAAAGTTATGAACCTGGGAGCTTTATTTCTT 2280  
QY  
2281 TCACATGCTTTTCTAATTTATGCTTGAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
Db  
2281 TCACATGCTTTTCTAATTTATGCTTGAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
QY  
2341 GCAAAAGCAGGGCCATTTATACCTGAACATAGAGAAGGAACACCCGTTTGTGTGCC 2400  
Db  
2341 GCAAAAGCAGGGCCATTTATACCTGAACATAGAGAAGGAACACCCGTTTGTGTGCC 2400  
QY  
2401 CCTGCTTGAAGGAAGAAATTAATCTCTGAAGTCTGGGCAACAGAGAAGCAATATGGAGCAGC 2460  
Db  
2401 CCTGCTTGAAGGAAGAAATTAATCTCTGAAGTCTGGGCAACAGAGAAGCAATATGGAGCAGC 2460  
QY  
2461 CAAAAGATGCGCGTCTGTTCAAGTTAACTAAAGGATTTCCACTTCTTCTTCCCTACCAA 2520  
Db  
2461 CAAAAGATGCGCGTCTGTTCAAGTTAACTAAAGGATTTCCACTTCTTCTTCCCTACCAA 2520  
QY  
2521 GGCAGTACCCCTCAGAGCCAGGCCCAACAGGATTTCCAAAAGATTTGTTAAGGACTTAA 2580  
Db  
2521 GGCAGTACCCCTCAGAGCCAGGCCCAACAGGATTTCCAAAAGATTTGTTAAGGACTTAA 2580  
QY  
2581 AAGCCCAAGGCTTAGTAAAAACCATGATAACTCCCTGCAAGTAAATCCGTTAGTGGATTGAG 2640  
Db  
2581 AAGCCCAAGGCTTAGTAAAAACCATGATAACTCCCTGCAAGTAAATCCGTTAGTGGATTGAG 2640



QY 2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAATCAATGG 2700  
DB |||||  
QY 2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAATCAATGG 2700  
DB |||||  
QY 2701 AGGCGGTGCTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATC 2760  
DB |||||  
QY 2701 AGGCGGTGCTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATC 2760  
DB |||||  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGACCTTACAGATGCTTCTGATCCCTG 2820  
DB |||||  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGACCTTACAGATGCTTCTGATCCCTG 2820  
DB |||||  
QY 2821 TACATCTGACTCTCAATCTCTGTTTGGCTTTCAGAGTACTTCAAAACCCARCATCTCAAC 2880  
DB |||||  
QY 2821 TACATCTGACTCTCAATCTCTGTTTGGCTTTCAGAGTACTTCAAAACCCARCATCTCAAC 2880  
DB |||||  
QY 2881 TCACCTGGACTTCTTACCCCAAGGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAT 2940  
DB |||||  
QY 2881 TCACCTGGACTTCTTACCCCAAGGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAT 2940  
DB |||||  
QY 2941 TAGCCCAAGACTTGAGYCARMTWCATACCTGGACACTCTTTGCTTCRGTAKGTGGATG 3000  
DB |||||  
QY 2941 TAGCCCAAGACTTGAGYCARMTWCATACCTGGACACTCTTTGCTTCRGTAKGTGGATG 3000  
DB |||||  
QY 3001 ATTACTTTTTRGCGCYRTTCAGAAACCTTGTGCGCATCAAGCCACCCAGRCCTCTTMA 3060  
DB |||||  
QY 3001 ATTACTTTTTRGCGCYRTTCAGAAACCTTGTGCGCATCAAGCCACCCAGRCCTCTTMA 3060  
DB |||||  
QY 3061 ATTCTCTCGYACCTGTGGCTTACAWGGTTTCCAAACCSARAGCTCARTCTGCTCACAGC 3120  
DB |||||  
QY 3061 ATTCTCTCGYACCTGTGGCTTACAWGGTTTCCAAACCSARAGCTCARTCTGCTCACAGC 3120  
DB |||||  
QY 3121 AGGTTAAATCTTAGRCCTAARATTAATCCAAAGGCACCCARGGCCCTCAGTGAGGAAYRYA 3180  
DB |||||  
QY 3121 AGGTTAAATCTTAGRCCTAARATTAATCCAAAGGCACCCARGGCCCTCAGTGAGGAAYRYA 3180  
DB |||||  
QY 3181 TCAGCCTTACTTGGCTTATCTTCATCYCAAAACCTTAAGCAACTAAGRRTTCTCTTG 3240  
DB |||||  
QY 3181 TCAGCCTTACTTGGCTTATCTTCATCYCAAAACCTTAAGCAACTAAGRRTTCTCTTG 3240  
DB |||||  
QY 3241 GRTAAAYAGGYTCTCCGAAWATGGATCCCCAGGTGGGCRANAATAGCCAGGYCATTA 3300  
DB |||||  
QY 3241 GRTAAAYAGGYTCTCCGAAWATGGATCCCCAGGTGGGCRANAATAGCCAGGYCATTA 3300  
DB |||||  
QY 3301 WATACASTAATTAAGGAACCTCAGAAAGCCAATACCCATTARTTAAGATGGAYAMCTGAA 3360  
DB |||||  
QY 3301 WATACASTAATTAAGGAACCTCAGAAAGCCAATACCCATTARTTAAGATGGAYAMCTGAA 3360  
DB |||||  
QY 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAAAGGCCCTTAAACCCCAAGYCCCAGTGTAAGYT 3420  
DB |||||  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTAYATRYTCAAGAAATAACAGAAAYAGCTCTRGA 3480  
DB |||||  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTAYATRYTCAAGAAATAACAGAAAYAGCTCTRGA 3480  
DB |||||  
QY 3481 GTCCCTTACAGRTCCRAGGAYGAGCTTGCACCCYRTGRCRYACCTGASTAGGAAYT 3540  
DB |||||  
QY 3481 GTCCCTTACAGRTCCRAGGAYGAGCTTGCACCCYRTGRCRYACCTGASTAGGAAYT 3540  
DB |||||  
QY 3541 GATGTAGTGCAAGGGTTGRCYCTCATTTTAYGGTGTGGTGGCAGTAGCAGYKTA 3600  
DB |||||  
QY 3541 GATGTAGTGCAAGGGTTGRCYCTCATTTTAYGGTGTGGTGGCAGTAGCAGYKTA 3600  
DB |||||  
QY 3601 GTATCTGAACAGCTTAAATAATACAGGGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB |||||  
QY 3601 GTATCTGAACAGCTTAAATAATACAGGGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB |||||  
QY 3661 GTGAAYRGCACTACTCTGCTTAAAGGAGCTTGTGGCTGTGACAGCAACYGTTTTACTTAAA 3720  
DB |||||  
QY 3661 GTGAAYRGCACTACTCTGCTTAAAGGAGCTTGTGGCTGTGACAGCAACYGTTTTACTTAAA 3720  
DB |||||

QY 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACCTTGTGCAACTCTTAAC 3780  
DB |||||  
QY 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACCTTGTGCAACTCTTAAC 3780  
DB |||||  
QY 3781 CCAGYCNCAATTTCTTCCAGACCAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
DB |||||  
QY 3781 CCAGYCNCAATTTCTTCCAGACCAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
DB |||||  
QY 3841 TCTCAAAACCTATGCGCACTCGAGGGACCTTCTTGTAGAGTTCCTTGTGACTGATCCYGCCTT 3900  
DB |||||  
QY 3841 TCTCAAAACCTATGCGCACTCGAGGGACCTTCTTGTAGAGTTCCTTGTGACTGATCCYGCCTT 3900  
DB |||||  
QY 3901 CCACTTGTATCTGATGGAAGTTCTTCTGTAGAAAGACCTTCGAAAGYGGGTATGC 3960  
DB |||||  
QY 3901 CCACTTGTATCTGATGGAAGTTCTTCTGTAGAAAGACCTTCGAAAGYGGGTATGC 3960  
DB |||||  
QY 3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACCTAGTCTTAA 4020  
DB |||||  
QY 3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACCTAGTCTTAA 4020  
DB |||||  
QY 4021 GCTRGCAAGACTTAATAGCCCTTCACTAGGGCACTAGAAATTTAGGAGAAAGAAAGGAAAA 4080  
DB |||||  
QY 4021 GCTRGCAAGACTTAATAGCCCTTCACTAGGGCACTAGAAATTTAGGAGAAAGAAAGGAAAA 4080  
DB |||||  
QY 4081 TATATATACAGACTCTTARTATCTTACCTAGTCTTCCATGCCATGMRGCAATATGSA 4140  
DB |||||  
QY 4081 TATATATACAGACTCTTARTATCTTACCTAGTCTTCCATGCCATGMRGCAATATGSA 4140  
DB |||||  
QY 4141 AGAAAGGGAAATTCCTAACTTCYAGRGGAACCTATCAMACATCAGGAAGCAATTAGGAR 4200  
DB |||||  
QY 4141 AGAAAGGGAAATTCCTAACTTCYAGRGGAACCTATCAMACATCAGGAAGCAATTAGGAR 4200  
DB |||||  
QY 4201 ATTATTATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGGGGTCTATCA 4260  
DB |||||  
QY 4201 ATTATTATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGGGGTCTATCA 4260  
DB |||||  
QY 4261 NAAAGAAAGRAAGGAAATASAGRGAAATYTGCCAGCAKATATTTGAAGCAAAAGAGC 4320  
DB |||||  
QY 4261 NAAAGAAAGRAAGGAAATASAGRGAAATYTGCCAGCAKATATTTGAAGCAAAAGAGC 4320  
DB |||||  
QY 4321 TCAGAGGAGGACCTTCCATTAAGAAATCTTAACTTCCTCTAGTATAGGTATCTC 4380  
DB |||||  
QY 4321 TCAGAGGAGGACCTTCCATTAAGAAATCTTAACTTCCTCTAGTATAGGTATCTC 4380  
DB |||||  
QY 4381 CTTTCGGGAAACCAAGCCCGCTTACCTCAGCAGAGAAACAGAAATGGGGAACCTCAGCAGG 4440  
DB |||||  
QY 4381 CTTTCGGGAAACCAAGCCCGCTTACCTCAGCAGAGAAACAGAAATGGGGAACCTCAGCAGG 4440  
DB |||||  
QY 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAGAGGGAATACTTTTGGCTGCAAC 4500  
DB |||||  
QY 4501 TATCCATGGAAATTTACTTAAACCCCTTCACTCAAACTTTCTAGGCAATGATAGCAC 4560  
DB |||||  
QY 4501 TATCCATGGAAATTTACTTAAACCCCTTCACTCAAACTTTCTAGGCAATGATAGCAC 4560  
DB |||||  
QY 4561 CCATCARATGGCCAAATCAATTTTCTGGAACCGCCCTTTTCAAACTATCAAGCARAT 4620  
DB |||||  
QY 4561 CCATCARATGGCCAAATCAATTTTCTGGAACCGCCCTTTTCAAACTATCAAGCARAT 4620  
DB |||||  
QY 4621 AKTCAGGGCTGTGAATGTGCAARAAATTAATCCCTGCTTATCGCAAGCTCTTTC 4680  
DB |||||  
QY 4621 AKTCAGGGCTGTGAATGTGCAARAAATTAATCCCTGCTTATCGCAAGCTCTTTC 4680  
DB |||||  
QY 4681 AGGAAACAAABAAACAGGCCATTACCTTGRABABACTGGCAACTGATTTTCCCAACAG 4740  
DB |||||  
QY 4681 AGGAAACAAABAAACAGGCCATTACCTTGRABABACTGGCAACTGATTTTCCCAACAG 4740  
DB |||||  
QY 4741 CCCAAACCTCAGGATTTTCACTTACTAGTCTGGTARATCTTTCACGGGTGGGCA 4800  
DB |||||  
QY 4741 CCCAAACCTCAGGATTTTCACTTACTAGTCTGGTARATCTTTCACGGGTGGGCA 4800  
DB |||||  
QY 4801 RAGGCTTCTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAA 4860  
DB |||||

[illegible]

5881	Db	GGAGTCTACTGTCTGTGGACATTACTTACCCCAAACTGGTATGTCTGATGGGGGTGGAGTT	594
5941	Qy	CAAGATCAGGCAAGAGAAAAACATGTAAAAAGAAAGTAACTCTCCAACTCACCSCGGGTACAT	600
5941	Db	CAAGATCAGGCNAGAGAAAAACATGTAAAAGAAAGTAACTCTCCAACTCACAACSCGGGTACAT	600
6001	Qy	GGCACTCTTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAAACCTCCGGTACC	606
6001	Db	GGCACTCTTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAAACCTCCGGTACC	606
6061	Qy	CATACTCGCTGGTGAAGCCTATTTAATACACACCTCACTGGGCTCCATGAGGCTTCGGCC	612
6061	Db	CATACTCGCTGGTGAAGCCTATTTAATACACACCTCACTGGGCTCCATGAGGCTTCGGCC	612
6121	Qy	CAAAACCTCTACTAACTGTGGATATGCTCCCTGAACTTTCARGCCCATATGTTTCAATC	618
6121	Db	CAAAACCTCTACTAACTGTGGATATGCTCCCTGAACTTTCARGCCCATATGTTTCAATC	618
6181	Qy	CCTGTACTGAAACAATGGAACAACTTCAGCACAGAAATAAACCACTTCCTGTTTATAGTA	624
6181	Db	CCTGTACTGAAACAATGGAACAACTTCAGCACAGAAATAAACCACTTCCTGTTTATAGTA	624
6241	Qy	GGACCTCTTGTTTCCCAATSTGGAATAAACCCATACCTCAAAACCTCAGCTGTGTAATAATTT	630
6241	Db	GGACCTCTTGTTTCCCAATSTGGAATAAACCCATACCTCAAAACCTCAGCTGTGTAATAATTT	630
6301	Qy	AGCAATACTACATACACAACAACCTCCCAATGATCAGGTGGGTAACTCCTCCACACAA	636
6301	Db	AGCAATACTACATACACAACAACCTCCCAATGATCAGGTGGGTAACTCCTCCACACAA	636
6361	Qy	ATAGTCTGCCTACCCCTCAGGAATAATTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG	642
6361	Db	ATAGTCTGCCTACCCCTCAGGAATAATTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG	642
6421	Qy	AATGGCTCTCAGAACTATGTGCTTCTCTCATTTTAGTGGCCCCCYATGRCACATCTAC	648
6421	Db	AATGGCTCTCAGAACTATGTGCTTCTCTCATTTTAGTGGCCCCCYATGRCACATCTAC	648
6481	Qy	ACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT	654
6481	Db	ACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT	654
6541	Qy	CTTCCTTTTGTATAGGACGAGGAGTCTAGTGTGCATCTAGGTACTTGGCATTTGGCGGTATC	660
6541	Db	CTTCCTTTTGTATAGGACGAGGAGTCTAGTGTGCATCTAGGTACTTGGCATTTGGCGGTATC	660
6601	Qy	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	666
6601	Db	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	666
6661	Qy	GTCCCGGACTCCCTGGTGCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCT	672
6661	Db	GTCCCGGACTCCCTGGTGCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCT	672
6721	Qy	CRAAATCGAAGAGCTTTAGACTYCTTAACCGCTGARAGAGGGGAACTGTTTATTTTTA	678
6721	Db	CRAAATCGAAGAGCTTTAGACTYCTTAACCGCTGARAGAGGGGAACTGTTTATTTTTA	678
6781	Qy	GGGGAAGAATGCTGTTATTTATGTTAAATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA	684
6781	Db	GGGGAAGAATGCTGTTATTTATGTTAAATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA	684
6841	Qy	ATTTCSAGATCGAATACACGTTAKAGCAGAGAGCTTCGAAACACTGACACCTCGGGCCTC	690
6841	Db	ATTTCSAGATCGAATACACGTTAKAGCAGAGAGCTTCGAAACACTGACACCTCGGGCCTC	690
6901	Qy	CTCAGCCCATGGATGGCCCTGGATTCTCCCTCTTTTAGGACCTCTTAGCAGCTATAATATTG	696
6901	Db	CTCAGCCCATGGATGGCCCTGGATTCTCCCTCTTTTAGGACCTCTTAGCAGCTATAATATTG	696
6961	Qy	CTACTCTCTTTTGGACCCCTGTATCTTTTAACTCTTGTATTAATTTGCTTCTTCAGAAATC	702
6961	Db	CTACTCTCTTTTGGACCCCTGTATCTTTTAACTCTTGTATTAATTTGCTTCTTCAGAAATC	702

QY 7021 GAAGCTGTRAACTACAAATGGAGCCCAAGATGCGAGTCCAAAGACTAAGATCTACCGGAGA 7080  
 DB |||||  
 QY 7021 GAAGCTGTRAACTACAAATGGAGCCCAAGATGCGAGTCCAAAGACTAAGATCTACCGGAGA 7080  
 DB |||||  
 QY 7081 CCCCTGGACCGGCTGTAGCCCAACGATCTGATGTTAATGACATCAAAAGGCACCCCTCCT 7140  
 DB |||||  
 QY 7081 CCCCTGGACCGGCTGTAGCCCAACGATCTGATGTTAATGACATCAAAAGGCACCCCTCCT 7140  
 DB |||||  
 QY 7141 GAGGAATCTCAGCTGCAAACTCTACTACGCCCAATTCAGAGGAAGCAAGTATAGAGC 7200  
 DB |||||  
 QY 7141 GAGGAATCTCAGCTGCAAACTCTACTACGCCCAATTCAGAGGAAGCAAGTATAGAGC 7200  
 DB |||||  
 QY 7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGGACTGAGA 7260  
 DB |||||  
 QY 7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGGACTGAGA 7260  
 DB |||||  
 QY 7261 GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCCTAGSTGGGAAGTG 7320  
 DB |||||  
 QY 7261 GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCCTAGSTGGGAAGTG 7320  
 DB |||||  
 QY 7321 ACCACATCCACCTTTAAACACCGGGCTTGCAACTTAGYTCAACCTGACCAATCAGAGAG 7380  
 DB |||||  
 QY 7321 ACCACATCCACCTTTAAACACCGGGCTTGCAACTTAGYTCAACCTGACCAATCAGAGAG 7380  
 DB |||||  
 QY 7381 CTCACATAAATGCTAATTAGGCAAGACAGGAGTAAAGAAATAGCCAATCATYATTGC 7440  
 DB |||||  
 QY 7381 CTCACATAAATGCTAATTAGGCAAGACAGGAGTAAAGAAATAGCCAATCATYATTGC 7440  
 DB |||||  
 QY 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCCCAAGTTCGAGCCGCGCAA 7500  
 DB |||||  
 QY 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCCCAAGTTCGAGCCGCGCAA 7500  
 DB |||||  
 QY 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
 DB |||||  
 QY 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
 DB |||||  
 QY 7561 TCTATTAAATCTTGCARCTGCR 7582  
 DB |||||  
 QY 7561 TCTATTAAATCTTGCARCTGCR 7582  
 DB |||||

RESULT 4  
 AX027480  
 LOCUS AX027480 7582 bp DNA linear PAT 16-SEP-2000  
 DEFINITION Sequence 30 from Patent FR2788784.  
 ACCESSION AX027480  
 VERSION AX027480.1 GI:10188444  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Mallet,F., Voisset,C. and Paranhos,B.G.  
 JOURNAL Patent: FR 2788784-A 30 28-JUL-2000;  
 BIO MERIEUX (FR)  
 FEATURES  
 Location/Qualifiers  
 1..7582  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 100.0%; Score 7582; DB 2; Length 7582;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCTTTGGGTCC 60  
 DB |||||  
 QY 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCTTTGGGTCC 60  
 DB |||||

QY 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCACACTGCA 120  
 DB |||||  
 QY 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCACACTGCA 120  
 DB |||||  
 QY 121 CTTCTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180  
 DB |||||  
 QY 121 CTTCTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180  
 DB |||||  
 QY 181 TGTTCGCAACCCGACACCTGCGCTGACTCCCATCCCTCTGGATCTTCTGAGGGTGC 240  
 DB |||||  
 QY 181 TGTTCGCAACCCGACACCTGCGCTGACTCCCATCCCTCTGGATCTTCTGAGGGTGC 240  
 DB |||||  
 QY 241 CGCTGCTCTGATCCAGGAGCGCCCAATTCGCGCTCCCAATTCGCGCTTAAAGGGCTGC 300  
 DB |||||  
 QY 241 CGCTGCTCTGATCCAGGAGCGCCCAATTCGCGCTCCCAATTCGCGCTTAAAGGGCTGC 300  
 DB |||||  
 QY 301 CATTTGTCCTGACGCTTAAGTCTGCTGGGTTTGTCTTAATTTGAGCTGAACTANTCACT 360  
 DB |||||  
 QY 301 CATTTGTCCTGACGCTTAAGTCTGCTGGGTTTGTCTTAATTTGAGCTGAACTANTCACT 360  
 DB |||||  
 QY 361 GGGTTCATGCTCTCTTCTGTGACCCACGCTTCTAATAKAACTATAACACTTACCACA 420  
 DB |||||  
 QY 361 GGGTTCATGCTCTCTTCTGTGACCCACGCTTCTAATAKAACTATAACACTTACCACA 420  
 DB |||||  
 QY 421 TGGCCCAAGATTCATCTCTTGGATCCGTGAGGSCAAAGAACTCCAGGTCAGAGATAC 480  
 DB |||||  
 QY 421 TGGCCCAAGATTCATCTCTTGGATCCGTGAGGSCAAAGAACTCCAGGTCAGAGATAC 480  
 DB |||||  
 QY 481 GARGCTTGGCACCCTCTTGGAGCGGCTGCTACCTCTTGGAGGCTTCCACCACTC 540  
 DB |||||  
 QY 481 GARGCTTGGCACCCTCTTGGAGCGGCTGCTACCTCTTGGAGGCTTCCACCACTC 540  
 DB |||||  
 QY 541 TTGGGAGCTCTGTGACCAAGGACCCCTCGGCTTACATTTTGGCRACCCAMSRACCGACATCC 600  
 DB |||||  
 QY 541 TTGGGAGCTCTGTGACCAAGGACCCCTCGGCTTACATTTTGGCRACCCAMSRACCGACATCC 600  
 DB |||||  
 QY 601 MAAGTATGGGAAAAGCTTCCCGCAAGACAAAACCGCCCTTAAGAGCTATTCTGGARAAT 660  
 DB |||||  
 QY 601 MAAGTATGGGAAAAGCTTCCCGCAAGACAAAACCGCCCTTAAGAGCTATTCTGGARAAT 660  
 DB |||||  
 QY 661 TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAAGCTTATATTTCTCTGACGTGC 720  
 DB |||||  
 QY 661 TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAAGCTTATATTTCTCTGACGTGC 720  
 DB |||||  
 QY 721 GCTTGGCACTCTGTGAGGAAAGTATAATTTAATACCACTTTACAGCTAGACTCTTTTG 780  
 DB |||||  
 QY 721 GCTTGGCACTCTGTGAGGAAAGTATAATTTAATACCACTTTACAGCTAGACTCTTTTG 780  
 DB |||||  
 QY 781 TAGAAAAGGCAATGGAGTGAAGTGCATTAAGTACAAACTTTTCTTCAATTAAGAGACAA 840  
 DB |||||  
 QY 781 TAGAAAAGGCAATGGAGTGAAGTGCATTAAGTACAAACTTTTCTTCAATTAAGAGACAA 840  
 DB |||||  
 QY 841 CTCACAAATTTATGAAAAGTGTATTTATGCGCTCAGAGAGCTTTCAGAGTCTACTCC 900  
 DB |||||  
 QY 841 CTCACAAATTTATGAAAAGTGTATTTATGCGCTCAGAGAGCTTTCAGAGTCTACTCC 900  
 DB |||||  
 QY 901 CTATCCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCCAATGG 960  
 DB |||||  
 QY 901 CTATCCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCCAATGG 960  
 DB |||||  
 QY 961 TCCAAAAGGAGATAGACAAAAGGTAACAGTGAAGAGTGCATATTTCCCAAT 1020  
 DB |||||  
 QY 961 TCCAAAAGGAGATAGACAAAAGGTAACAGTGAAGAGTGCATATTTCCCAAT 1020  
 DB |||||  
 QY 1021 TATGACCCCTCCCAAGCAGTGGAGAGAGAAATTCGCGCCAGCGAGTGCATGTGCT 1080  
 DB |||||  
 QY 1021 TATGACCCCTCCCAAGCAGTGGAGAGAGAAATTCGCGCCAGCGAGTGCATGTGCT 1080  
 DB |||||  
 QY 1081 TTTTCTCTCCAGACTTAAAGCAATATAAAGAGTAAATTTCTCAGATAATCCCT 1140  
 DB |||||  
 QY 1081 TTTTCTCTCCAGACTTAAAGCAATATAAAGAGTAAATTTCTCAGATAATCCCT 1140  
 DB |||||  
 QY 1141 GATGGCTATATTTGRTGTTTACAAAGGTTAGGACAAATTTCTTGTATCTGACATGGAGAT 1200

1141	Db	 GATGGCTATATTGRTGTTTTTACAAGGGTTAGGCAATTTCTTTGATCTGACATCGAGAGAT	1200
1201	Qy	ATATATGTCCTACTGCTAAATCAGACACATTAACCCCAATGAGAGAAGTGCCACCAATAACTGCG	1260
1201	Db	ATATATGTCCTACTGCTAAATCAGACACATTAACCCCAATGAGAGAAGTGCCACCAATAACTGCG	1260
1261	Qy	AGCCTGAGRGTTTTGGCGGATCTCTGGTATCTCAGTCAGGTCNAATGGATPANGGATGCAACA	1320
1261	Db	AGCCTGAGRGTTTTGGCGGATCTCTGGTATCTCAGTCAGGTCNAATGGATPANGGATGCAACA	1320
1321	Qy	GAAGGAAGANAATGATTTCCCAACAGGCCAGCARGCAGTTCACAGTCTASACCTCATTTG	1380
1321	Db	GAAGGAAGANAATGATTTCCCAACAGGCCAGCARGCAGTTCACAGTCTASACCTCATTTG	1380
1381	Qy	GGGACACAGAAATCAGTAACATCGGAGATTTGGTCTGCACACATTTTGCTAACTTGTGTGC	1440
1381	Db	GGGACACAGAAATCAGTAACATCGGAGATTTGGTCTGCACACATTTTGCTAACTTGTGTGC	1440
1441	Qy	TASAAGGACTAAGGAAAACCTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATTA	1500
1441	Db	TASAAGGACTAAGGAAAACCTASGAAGAAARTCTAYGAATTACTCAATGATGTCCACCATTA	1500
1501	Qy	ACACAGGGGAAGGGAAGAAATCTCTA CTGCTTTTCTGGAGAGACTAAGGAGGCGATTGAG	1560
1501	Db	ACACAGGGGAAGGGAAGAAATCTCTA CTGCTTTTCTGGAGAGACTAAGGAGGCGATTGAG	1560
1561	Qy	GAAGCGTGCCCTCTCTGTCACCTGACTCTTCTGAAGGCCAATAATCTTAAGCGTAAAGTT	1620
1561	Db	GAAGCGTGCCCTCTCTGTCACCTGACTCTTCTGAAGGCCAATAATCTTAAGCGTAAAGTT	1620
1621	Qy	TATCACTCAGTCAGCTGCAGACATTAAGAAAAAATCTCAAAAGCTGCGCTAGGCCCGGAG	1680
1621	Db	TATCACTCAGTCAGCTGCAGACATTAAGAAAAAATCTCAAAAGCTGCGCTAGGCCCGGAG	1680
1681	Qy	CAAAACTTTAGAAACCTTATTGAACTTTGGCAA CYTCGGTTTTTTTATAATAGAGATCAGGAG	1740
1681	Db	CAAAACTTTAGAAACCTTATTGAACTTTGGCAA CYTCGGTTTTTTTATAATAGAGATCAGGAG	1740
1741	Qy	GAGCAGGGGNAACAGGACAAACGGGATTAATAAAAGGCCACCGCTTTAGTCATGACCCT	1800
1741	Db	GAGCAGGGGNAACAGGACAAACGGGATTAATAAAAGGCCACCGCTTTAGTCATGACCCT	1800
1801	Qy	CAGGCAAGTGACATTTGGAGGCTCTGAAAGGGAAAAAGCTGGGCAAAATGMAATGCCTAA	1860
1801	Db	CAGGCAAGTGACATTTGGAGGCTCTGAAAGGGAAAAAGCTGGGCAAAATGMAATGCCTAA	1860
1861	Qy	TAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAGATTTGTCCTCAAGTAGAAG	1920
1861	Db	TAGGGCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAGATTTGTCCTCAAGTAGAAG	1920
1921	Qy	TAAAGCGGCCCTTCGTGATGCCCTTTATTTCAAGGGAATCAGTCGGAAGGCCACCTGCC	1980
1921	Db	TAAAGCGGCCCTTCGTGATGCCCTTTATTTCAAGGGAATCAGTCGGAAGGCCACCTGCC	1980
1981	Qy	CAGGGGACAAAGGCTTTTGTAGTCAGAGCCACTAACAGATGATCAGCAGCAGGAGACTG	2040
1981	Db	CAGGGGACAAAGGCTTTTGTAGTCAGAGCCACTAACAGATGATCAGCAGCAGGAGACTG	2040
2041	Qy	AGGGTGCTCGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTT	2100
2041	Db	AGGGTGCTCGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTT	2100
2101	Qy	GACATTTAGGGCCAGGAAGTTGTCTCTGGAACACTGGTGCGGTCTTTCTTAGCTTTACT	2160
2101	Db	GACATTTAGGGCCAGGAAGTTGTCTCTGGAACACTGGTGCGGTCTTTCTTAGCTTTACT	2160
2161	Qy	CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACG	2220
2161	Db	CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACG	2220
2221	Qy	GGCAGTCACTAGATACTTTTTCAGGCCACTAAAGTTATGAACCTGGGAGCTTTATTCTTTT	2280

Dd		2221	GGCAGTCACTAGANACTTTTCCACGCGCACTAAGTTATGAACGTGGGAGCTTTATTCTCTT	2280
Qy		2281	TCACTGCTTTTTCTAAATTATGCTTGAAAGCCCCACACTACCTTGTGTAGGGAGAGACAATTCTA	2340
Dd		2281	TCACATGCTTTTCTAANTATGCTTTGAAAGCCCCACACTACCTTGTGTAGGGAGAGACAATTCTA	2340
Qy		2341	GCAAAAGCAGGGGCCAATTATACACCTGAAACATAGGAGAAGGAACAACCGTTTGTGTGNCC	2400
Dd		2341	GCAAAAGCAGGGGCCAATTATACACCTGAAACATAGGAGAAGGAACAACCGTTTGTGTGNCC	2400
Qy		2401	CCTGCTTGAGGAGGAATAATCCTCAAGTCTCTGGGCACAACAGAAGGACAATATGACGAGC	2460
Dd		2401	CCTGCTTGAGGAGGAATAATCCTCAAGTCTCTGGGCACAACAGAAGGACAATATGACGAGC	2460
Qy		2461	CAAAGAATGCCGTCCTGTTTCAAGTTTAAACTAAAGGATTCACATCTCTTTCCCCTACCAAA	2520
Dd		2461	CAAAGAATGCCGTCCTGTTTCAAGTTTAAACTAAAGGATTCACATCTCTTTCCCCTACCAAA	2520
Qy		2521	GGCAGTACCCCTCAGACCCAAGGCCCAACAAGGATTCAAAAAGATTTGTTTAAGGACTTAA	2580
Dd		2521	GGCAGTACCCCTCAGACCCAAGGCCCAACAAGGATTCAAAAAGATTTGTTTAAGGACTTAA	2580
Qy		2581	AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAGTAATTCGGTAGTGGATTGAG	2640
Dd		2581	AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAGTAATTCGGTAGTGGATTGAG	2640
Qy		2641	GAGGCAAGAAAACCCAGTGGACAGTGGAGGTTAGTGCAGATCTCAGGAATTAATCAATGG	2700
Dd		2641	GAGGCAAGAAAACCCAGTGGACAGTGGAGGTTAGTGCAGATCTCAGGAATTAATCAATGG	2700
Qy		2701	AGGCCGTTGCTTTTATACCCAGCTGTACTAGCCCTTATACTGTGMVTTCCCAATAC	2760
Dd		2701	AGGCCGTTGCTTTTATACCCAGCTGTACTAGCCCTTATACTGTGMVTTCCCAATAC	2760
Qy		2761	CAGAGAACAGAGTGGTTTTACSTCCTGAGACCTTWAGGATGCTTCTTCGTGATCCCTG	2820
Dd		2761	CAGAGAACAGAGTGGTTTTACSTCCTGAGACCTTWAGGATGCTTCTTCGTGATCCCTG	2820
Qy		2821	TACATCCTGACTCTCAATCTCTGTTTGCTTTTCAAGATCTTCAAAACCCARCATCTCAAC	2880
Dd		2821	TACATCCTGACTCTCAATCTCTGTTTGCTTTTGAAGATCTTCAAAACCCARCATCTCAAC	2880
Qy		2881	TCACCTGGACTRTTTATCCCAAGGGTTCAGGGATAGYCCCACTATNTTTGGCCAGGCAAT	2940
Dd		2881	TCACCTGGACTRTTTATCCCAAGGGTTCAGGGATAGYCCCACTATNTTTGGCCAGGCAAT	2940
Qy		2941	TAGCCCAAGACTTGAGYCARTYMTCACTGAGACACTTGTGCTTCRGTAKGTGGATG	3000
Dd		2941	TAGCCCAAGACTTGAGYCARTYMTCACTGAGACACTTGTGCTTCRGTAKGTGGATG	3000
Qy		3001	ATTTACTTTTRGCVGCYRPTTCAGAAACCTTGTGCCATCAAGCACCCACAGRCCTCTTMA	3060
Dd		3001	ATTTACTTTTRGCVGCYRPTTCAGAAACCTTGTGCCATCAAGCACCCACAGRCCTCTTMA	3060
Qy		3061	ATTTCTTCGYACCTGTGGCTACAWGGTTTCCAAACSARAGCTCARCTCTGCTCACAGC	3120
Dd		3061	ATTTCTTCGYACCTGTGGCTACAWGGTTTCCAAACSARAGCTCARCTCTGCTCACAGC	3120
Qy		3121	AGGTTAAATACTTAGGRCTAARATATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYA	3180
Dd		3121	AGGTTAAATACTTAGGRCTAARATATCCAAAGGCACCARGGCCCTCAGTGAGGAAYRYA	3180
Qy		3181	TCCAGCCTATCTGGCTTATCCTCANTCYCHAAACCTTAAGCACTAAGRRTTCCCTWG	3240
Dd		3181	TCCAGCCTATCTGGCTTATCCTCANTCYCHAAACCTTAAGCACTAAGRRTTCCCTWG	3240
Qy		3241	GCRTAAYAGGYTCTGCGCAAWATGGATTCCCCAGGTTWTCGCAAAATAGCCAGGYCATTA	3300
Dd		3241	GCRTAAYAGGYTCTGCGCAAWATGGATTCCCCAGGTTWTCGCAAAATAGCCAGGYCATTA	3300
Qy		3301	WATACASTAATAAGGAAACTCAGAAAGCCAATACCCATTTARTPAAGATGGAYAMCTGAA	3360
Dd		3301	WATACASTAATAAGGAAACTCAGAAAGCCAATACCCATTTARTPAAGATGGAYAMCTGAA	3360

QY 3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCAAGYCCCAGTGTAAAGYT 3420  
DB |||||  
QY 3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCAAGYCCCAGTGTAAAGYT 3420  
DB |||||  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTYATAYRTACAGAAAAAAGCAAGAAVAGCTCTRGGA 3480  
DB |||||  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTYATAYRTACAGAAAAAAGCAAGAAVAGCTCTRGGA 3480  
DB |||||  
QY 3481 GTCTTACACAGRTCCRAGGAGVAGCTTGGACCCYRTGGCRVACCTGASTAAGGAAAYT 3540  
DB |||||  
QY 3481 GTCTTACACAGRTCCRAGGAGVAGCTTGGACCCYRTGGCRVACCTGASTAAGGAAAYT 3540  
DB |||||  
QY 3541 GATGTAGTGGCAAGGGTTCRCYTCATGTTTAYGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
DB |||||  
QY 3541 GATGTAGTGGCAAGGGTTCRCYTCATGTTTAYGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
DB |||||  
QY 3601 GTATCTGAAGCAGTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB |||||  
QY 3601 GTATCTGAAGCAGTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB |||||  
QY 3661 GTGAAYRGCACTACTGCTAAAGGAGACTTGTGGCTGTGAGCAACACYGTTTACTTAAA 3720  
DB |||||  
QY 3661 GTGAAYRGCACTACTGCTAAAGGAGACTTGTGGCTGTGAGCAACACYGTTTACTTAAA 3720  
DB |||||  
QY 3721 TRTCAGGCTCTATTACTTGAAGGCCAGTCTGCRACCTGTGCACTTGTGCACTCTTAAAC 3780  
DB |||||  
QY 3721 TRTCAGGCTCTATTACTTGAAGGCCAGTCTGCRACCTGTGCACTTGTGCACTCTTAAAC 3780  
DB |||||  
QY 3781 CCAGYCNCACTTCTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
DB |||||  
QY 3781 CCAGYCNCACTTCTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
DB |||||  
QY 3841 TCTCAAACTATGCCACTCGAGGGGACCTTCTTGTAGAGTTCCTTGTGACTGATCCYGCCTT 3900  
DB |||||  
QY 3841 TCTCAAACTATGCCACTCGAGGGGACCTTCTTGTAGAGTTCCTTGTGACTGATCCYGCCTT 3900  
DB |||||  
QY 3901 CAACCTTGATCTAGTGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
DB |||||  
QY 3901 CAACCTTGATCTAGTGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
DB |||||  
QY 3961 AGTGGTCAGTGATAATAGGGAATATTTGAAAGTAATCCCTCACCTCCAGGAACTAGTCTYA 4020  
DB |||||  
QY 3961 AGTGGTCAGTGATAATAGGGAATATTTGAAAGTAATCCCTCACCTCCAGGAACTAGTCTYA 4020  
DB |||||  
QY 4021 GCTRGCAAGCACTAATAGCCYTCAYTRGGGCACTAGAAATAGGAGAAAGRAAAAGGYAAA 4080  
DB |||||  
QY 4021 GCTRGCAAGCACTAATAGCCYTCAYTRGGGCACTAGAAATAGGAGAAAGRAAAAGGYAAA 4080  
DB |||||  
QY 4081 TATATATACAGACTCTTRATPATCTYACCTAGTCTNTCCATGCCCATGMRGCAATATGSAR 4140  
DB |||||  
QY 4081 TATATATACAGACTCTTRATPATCTYACCTAGTCTNTCCATGCCCATGMRGCAATATGSAR 4140  
DB |||||  
QY 4141 AGAAAGGGAATTCCTTAATCTCYAGRGACACCTATCNACATCAGGAAGCCATTAGGAR 4200  
DB |||||  
QY 4141 AGAAAGGGAATTCCTTAATCTCYAGRGACACCTATCNACATCAGGAAGCCATTAGGAR 4200  
DB |||||  
QY 4201 ATTATATTYGGGWTACAGAAACCTARAGAGTGGMAGTCTTACACTGCYGGGGTCAATCA 4260  
DB |||||  
QY 4201 ATTATATTYGGGWTACAGAAACCTARAGAGTGGMAGTCTTACACTGCYGGGGTCAATCA 4260  
DB |||||  
QY 4261 NAAAGGAAAGRAAAGGGAATAASAAGRGAAATGCCAAGCAKATATTTCAAGCMAAAAGAGC 4320  
DB |||||  
QY 4261 NAAAGGAAAGRAAAGGGAATAASAAGRGAAATGCCAAGCAKATATTTCAAGCMAAAAGAGC 4320  
DB |||||  
QY 4321 TGCAAGGCAAGGACCTCCATTAGAAATGCTTTAACTTCCCTTATAGTATAGGGTAATCC 4380  
DB |||||  
QY 4321 TGCAAGGCAAGGACCTCCATTAGAAATGCTTTAACTTCCCTTATAGTATAGGGTAATCC 4380  
DB |||||  
QY 4381 CTTCCGGGAAACCAAGCCAGTACTCAGCAGGAGAAACAGATGGGAACTCAGAGG 4440  
DB |||||  
QY 4381 CTTCCGGGAAACCAAGCCAGTACTCAGCAGGAGAAACAGATGGGAACTCAGAGG 4440  
DB |||||

QY 4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGAAAAATATCTTTTGCCTGCAAC 4500  
DB |||||  
QY 4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGAAAAATATCTTTTGCCTGCAAC 4500  
DB |||||  
QY 4501 TATCCAATGAAATTAATTAATAAACCTTCAATCAAACTTTTCACTTAGGCATCGATAGCAC 4560  
DB |||||  
QY 4501 TATCCAATGAAATTAATTAATAAACCTTCAATCAAACTTTTCACTTAGGCATCGATAGCAC 4560  
DB |||||  
QY 4561 CCATCARATGGCCAAATCAATTAATTTACTGGACAGGCTTTTCAAACTATCAAGCARAT 4620  
DB |||||  
QY 4561 CCATCARATGGCCAAATCAATTAATTTACTGGACAGGCTTTTCAAACTATCAAGCARAT 4620  
DB |||||  
QY 4621 AKTCAGGGCTGTGAATGTGCCBARAAATAATCCCTGCCCTTATCGCCAAAGCTCTTC 4680  
DB |||||  
QY 4621 AKTCAGGGCTGTGAATGTGCCBARAAATAATCCCTGCCCTTATCGCCAAAGCTCTTC 4680  
DB |||||  
QY 4681 AGGARAACAAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTATCCCAAG 4740  
DB |||||  
QY 4681 AGGARAACAAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTATCCCAAG 4740  
DB |||||  
QY 4741 CCCAAACCTCAGGGATTTTCACTAGTCTGGGTARATATCTTTCAOGGTTGGGCA 4800  
DB |||||  
QY 4741 CCCAAACCTCAGGGATTTTCACTAGTCTGGGTARATATCTTTCAOGGTTGGGCA 4800  
DB |||||  
QY 4801 RAGGCCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTATAAAGGCACTAGTTTCATGAA 4860  
DB |||||  
QY 4801 RAGGCCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTATAAAGGCACTAGTTTCATGAA 4860  
DB |||||  
QY 4861 ATAAATTTCCAGATTTCCGAGCTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTCCAG 4920  
DB |||||  
QY 4861 ATAAATTTCCAGATTTCCGAGCTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTCCAG 4920  
DB |||||  
QY 4921 GCCACAGTAACCCAGGGAGTATCCAGGGGTTAGGTATAGCATATCACTTACACTGGCC 4980  
DB |||||  
QY 4921 GCCACAGTAACCCAGGGAGTATCCAGGGGTTAGGTATAGCATATCACTTACACTGGCC 4980  
DB |||||  
QY 4981 TGAAGGCCACAGTCTCTCAGGGAAGTTCGAGAAATGAATGAAYATCAAAAGACATCTA 5040  
DB |||||  
QY 4981 TGAAGGCCACAGTCTCTCAGGGAAGTTCGAGAAATGAATGAAYATCAAAAGACATCTA 5040  
DB |||||  
QY 5041 AAAAGCAAAACCCAGGAACCCACCTCACATGGCTGTCTGTGGCTATAGCCTTAAAA 5100  
DB |||||  
QY 5041 AAAAGCAAAACCCAGGAACCCACCTCACATGGCTGTCTGTGGCTATAGCCTTAAAA 5100  
DB |||||  
QY 5101 AGAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAGG 5160  
DB |||||  
QY 5101 AGAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAGG 5160  
DB |||||  
QY 5161 CCTTTCATAAACCAATGACCTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCA 5220  
DB |||||  
QY 5161 CCTTTCATAAACCAATGACCTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCA 5220  
DB |||||  
QY 5221 CCTCTTAGCCAAATATCAAGATCTTTAAACATTACAGGAACCTATCCCTGAGAAG 5280  
DB |||||  
QY 5221 CCTCTTAGCCAAATATCAAGATCTTTAAACATTACAGGAACCTATCCCTGAGAAG 5280  
DB |||||  
QY 5281 AGGAAAAAGAACTATTTCACCCMWGTGACATGATGATTTAGTCAAGTCCCTCTCTAATT 5340  
DB |||||  
QY 5281 AGGAAAAAGAACTATTTCACCCMWGTGACATGATGATTTAGTCAAGTCCCTCTCTAATT 5340  
DB |||||  
QY 5341 CCCCCTCCCTAGATACATCTCGGAGGACCCCTACCCAGTCAATTTTATYATACCCCACTG 5400  
DB |||||  
QY 5341 CCCCCTCCCTAGATACATCTCGGAGGACCCCTACCCAGTCAATTTTATYATACCCCACTG 5400  
DB |||||  
QY 5401 CGGTTAAAGTGGCTGAGTGGAGTCTTGGATATCATCACTTGAGTCAAAATCTTGATAC 5460  
DB |||||  
QY 5401 CGGTTAAAGTGGCTGAGTGGAGTCTTGGATATCATCACTTGAGTCAAAATCTTGATAC 5460  
DB |||||  
QY 5461 TGCCAAAGGAACTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGG 5520  
DB |||||  
QY 5461 TGCCAAAGGAACTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGG 5520  
DB |||||  
QY 5521 ATTTGCGCTGCTCTTCAAAACCAACAGGAGGAAAGTAACTAAAAATCATAAATCCCCC 5580  
DB |||||

|||||  
5521 ATTTCGGCTGCTCTTCAACAAACACGAGGAGAACTAACTAAATCAATAATCCCCC 5580  
|||||  
5581 ATGSGCCCTCCCTTATCAATAATTTTTCTCTKTAAGTSTTTTTYACCTSTTTTCACTCTCACT 5640  
|||||  
5581 ATGSGCCCTCCCTTATCAATAATTTTTCTCTKTAAGTSTTTTTYACCTSTTTTCACTCTCACT 5640  
|||||  
5641 GCACCCCTCCATGCGCTGTATGACCACTAGTACCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
|||||  
5641 GCACCCCTCCATGCGCTGTATGACCACTAGTACCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
|||||  
5701 ATGACAGCGTCCCGAAATATTCATGCCCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCC 5760  
|||||  
5701 ATGACAGCGTCCCGAAATATTCATGCCCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCC 5760  
|||||  
5761 ACCTTCACGTGCCACACCCATATGCGCGCAACTGCTATCACTCTGCGCACTCTTTTGCAATG 5820  
|||||  
5761 ACCTTCACGTGCCACACCCATATGCGCGCAACTGCTATCACTCTGCGCACTCTTTTGCAATG 5820  
|||||  
5821 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGACTT 5880  
|||||  
5821 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGACTT 5880  
|||||  
5881 GGAGTCACGTCTGTTGGACTTACTTACCCCAAACTGGTATCTCTGATGGGGGTGGAGTT 5940  
|||||  
5881 GGAGTCACGTCTGTTGGACTTACTTACCCCAAACTGGTATCTCTGATGGGGGTGGAGTT 5940  
|||||  
5941 CAAGATCAGGCAAGAAAAAATGTAAGAAAGTAATCTCCCACTCAACCCGGGTACAT 6000  
|||||  
5941 CAAGATCAGGCAAGAAAAAATGTAAGAAAGTAATCTCCCACTCAACCCGGGTACAT 6000  
|||||  
6001 GGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAACTACATGAATCCCTCCGTACC 6060  
|||||  
6001 GGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAACTACATGAATCCCTCCGTACC 6060  
|||||  
6061 CATACTCGCTGCTGTAAGCTTATTAATACACCCCTCAGTGGCTCCATGAGTCTCGGCC 6120  
|||||  
6061 CATACTCGCTGCTGTAAGCTTATTAATACACCCCTCAGTGGCTCCATGAGTCTCGGCC 6120  
|||||  
6121 CAAACCCCTACTAACTGTTGGATATGCTTCCCTCCCTGAACTTCARGGCAATGTTTCAATC 6180  
|||||  
6121 CAAACCCCTACTAACTGTTGGATATGCTTCCCTCCCTGAACTTCARGGCAATGTTTCAATC 6180  
|||||  
6181 CTTGTAACCTGAAATGGAACAACTTCAGCACAGAAATAAACAACCACTTCCTGTTTGTAGTA 6240  
|||||  
6181 CTTGTAACCTGAAATGGAACAACTTCAGCACAGAAATAAACAACCACTTCCTGTTTGTAGTA 6240  
|||||  
6241 GGACCTCTGTTTCCCAATSTGGAATAACCCATACCTCAAACTCAACCTGTGTAAATTT 6300  
|||||  
6241 GGACCTCTGTTTCCCAATSTGGAATAACCCATACCTCAAACTCAACCTGTGTAAATTT 6300  
|||||  
6301 AGCAATACTACATACACAACTCCCAATGATCAGGTGGGTAACTCCTCCCAACAA 6360  
|||||  
6301 AGCAATACTACATACACAACTCCCAATGATCAGGTGGGTAACTCCTCCCAACAA 6360  
|||||  
6361 ATAGTCTGCTACCTCAGAAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTG 6420  
|||||  
6361 ATAGTCTGCTACCTCAGAAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTG 6420  
|||||  
6421 AATGGCTCTTCAGAATCTATGTCCTCTCATTTAGTGGCCCTTATGRCATCTAC 6480  
|||||  
6421 AATGGCTCTTCAGAATCTATGTCCTCTCATTTAGTGGCCCTTATGRCATCTAC 6480  
|||||  
6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
|||||  
6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
|||||  
6541 CTTCCCTTTGTTATAGACAGAGTCTAGTGGCACTAGTGGCAITGGCGGTATC 6600  
|||||  
6541 CTTCCCTTTGTTATAGACAGAGTCTAGTGGCACTAGTGGCAITGGCGGTATC 6600  
|||||  
6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACG 6660  
|||||

Db  
6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGAATGGAACG 6660  
Qy  
6661 GTGCGGACTCCCTGGTCACTTGGTCAAGATCAACTTAATCTCCCTAGCAGCAGTACTCTT 6720  
Db  
6661 GTGCGGACTCCCTGGTCACTTGGTCAAGATCAACTTAATCTCCCTAGCAGCAGTACTCTT 6720  
Qy  
6721 CRAAATCGAAGAGCTTTAGACTYGTAACTCCCTGARAGAGGGGAACTCTGTTTATTTT 6780  
Db  
6721 CRAAATCGAAGAGCTTTAGACTYGTAACTCCCTGARAGAGGGGAACTCTGTTTATTTT 6780  
Qy  
6781 GGGGAAGAATCTGTTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Db  
6781 GGGGAAGAATCTGTTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Qy  
6841 ATTCSAGATCGAATACAACTGTAAGAGAGCTTCCGAAACATCGACCTGGGCGCTC 6900  
Db  
6841 ATTCSAGATCGAATACAACTGTAAGAGAGCTTCCGAAACATCGACCTGGGCGCTC 6900  
Qy  
6901 CTCAGCCRAATGGATGCGCTGGATCTCCCTCTTCTAGGACCTCTAGCAGCTATATATTTG 6960  
Db  
6901 CTCAGCCRAATGGATGCGCTGGATCTCCCTCTTCTAGGACCTCTAGCAGCTATATATTTG 6960  
Qy  
6961 CTACTCTCTTTTGGACCTGTATCTTTTACCTCTCTTGTAACTTTGTCTCTTCCAGAACT 7020  
Db  
6961 CTACTCTCTTTTGGACCTGTATCTTTTACCTCTCTTGTAACTTTGTCTCTTCCAGAACT 7020  
Qy  
7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATCGAGTCCAAGACTTAAGTCTACCGCAGA 7080  
Db  
7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATCGAGTCCAAGACTTAAGTCTACCGCAGA 7080  
Qy  
7081 CCCCTGGACCGCTGTAGCCCAAGATCTGATGTTTAAATGATCAATCAAAAGGCAACCTCTCT 7140  
Db  
7081 CCCCTGGACCGCTGTAGCCCAAGATCTGATGTTTAAATGATCAATCAAAAGGCAACCTCTCT 7140  
Qy  
7141 GAGGAAATCTCAGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
Db  
7141 GAGGAAATCTCAGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
Qy  
7201 GGTSGTGGGCAACCTCCCGCAACAGCCTTAGTGTCTCTGTTGAGTGGGGACTGAGA 7260  
Db  
7201 GGTSGTGGGCAACCTCCCGCAACAGCCTTAGTGTCTCTGTTGAGTGGGGACTGAGA 7260  
Qy  
7261 GACGAGCTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGTGGGAGGTTG 7320  
Db  
7261 GACGAGCTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGTGGGAGGTTG 7320  
Qy  
7321 ACCCATCCACTTTTAAACACGGGCTTGCATCTTAGTGTCACTGACCAATCAGAGAG 7380  
Db  
7321 ACCCATCCACTTTTAAACACGGGCTTGCATCTTAGTGTCACTGACCAATCAGAGAG 7380  
Qy  
7381 CTCACATAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCATCATTTATTCG 7440  
Db  
7381 CTCACATAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCATCATTTATTCG 7440  
Qy  
7441 MTGAGACACAGCAGGAGGCAATGATCGGATATAAACCCCAAGTTCGAGCCGGCAA 7500  
Db  
7441 MTGAGACACAGCAGGAGGCAATGATCGGATATAAACCCCAAGTTCGAGCCGGCAA 7500  
Qy  
7501 CGGCAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTCAATGCTATTTTAC 7560  
Db  
7501 CGGCAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTCAATGCTATTTTAC 7560  
Qy  
7561 TCTATTAATCTTGCARCTGCR 7582  
Db  
7561 TCTATTAATCTTGCARCTGCR 7582

RESULT 5  
BD196244  
LOCUS  
DEFINITION  
ACCESSION

BD196244  
2372 bp  
DNA  
linear  
PAT 17-JUL-2003  
Endogenous retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.  
BD196244



```

VERSION      BD196244.1 GI:33006014
KEYWORDS     JP 2002512530-A/10.
SOURCE       unidentified
ORGANISM      unclassified sequences.
REFERENCE     1 (bases 1 to 2372)
AUTHORS       Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
TITLE         Endogenous retroviral sequences, associated with autoimmune
              diseases and/or with pregnancy disorders
JOURNAL       Patent: JP 2002512530-A 10 23-APR-2002;
COMMENT       BIO MERIEUX
              OS Unidentified
              PN JP 2002512530-A/10
              PD 23-APR-2002
              PE 06-JUL-1998 JP 1999508244
              PR 07-JUL-1997 FR 97/08815
              PI FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
              MANDRAND,
              PI FRANCOIS MALLET
              PC C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC
              CC Strandedness: Single;
              CC Topology: Linear;
              CC Endogenous retroviral sequences, associated with autoimmune
              diseases
              CC and/or with pregnancy disorders
              FH Key Location/Qualifiers
              FT source 1..2372
              FT Location/Qualifiers
              1..2372
              /organism="Unidentified".
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"

ORIGIN
Query Match 11.7%; Score 888; DB 2; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGGAGGGAGAAATCTACTGCTTCTCGAGAGACTAAGGAGGCGATTGAGGAAGCG 1566
DB 1383 GGGAGGGAGAAATCTACTGCTTCTCGAGAGACTAAGGAGGCGATTGAGGAAGCG 1442
QY 1567 TGCCTCTCTGACCTGACTCTTCTGAAGCCCACTAACTCTTAAAGCGTAAGTTATCAC 1626
DB 1443 TGCCTCTCTGACCTGACTCTTCTGAAGCCCACTAACTCTTAAAGCGTAAGTTATCAC 1502
QY 1627 TCAGTCAGCTGCAGACATTAGAAAACCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAAC 1686
DB 1503 TCAGTCAGCTGCAGACATTAGAAAACCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAAC 1562
QY 1687 TTAGAAACCTTATTGAACCTTGGCACTCGGTTTATTAATAGAGATCAGGAGGAGCAG 1746
DB 1563 TTAGAAACCTTATTGAACCTTGGCACTCGGTTTATTAATAGAGATCAGGAGGAGCAG 1622
QY 1747 GCGGAACAGACAACACGGGATTAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1806
DB 1623 GCGGAACAGACAACACGGGATTAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1682
QY 1807 AGTGGACTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAAAATTGATCCCTAATAGGC 1866
DB 1683 AGTGGACTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAAAATTGATCCCTAATAGGC 1742
QY 1867 TTGCTTCCAGTCGGGTCTACAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCC 1926
DB 1743 TTGCTTCCAGTCGGGTCTACAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCC 1802
QY 1927 GCCCTCTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCGGAGGG 1986
DB 1803 GCCCTCTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCGGAGGG 1862
QY 1987 ACAAGGCTCTTTAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGACTGAGGGTG 2046

```

```

Db 1863 ACAAGGCTCTTTAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGAGCTGAGGGTG 1922
QY 2047 CCTGGGCAAGGCCCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTGACCAT 2106
Db 1923 CCTGGGCAAGGCCCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTGACCAT 1982
QY 2107 TGAGGGCCAGGAAGGTGCTCTCTGACACATGCTGGTGGGCTCTTTAGTCTTACTCTTCTG 2166
Db 1983 TGAGGGCCAGGAAGGTGCTCTCTGACACATGCTGGTGGGCTCTTTAGTCTTACTCTTCTG 2042
QY 2167 TCCCGGACAACTGCTCTCCAGATCTCTCACTATTCTGAGGGGTCCTNTAAGACGGCAGT 2226
Db 2043 TCCCGGACAACTGCTCTCCAGATCTCTCACTATTCTGAGGGGTCCTNTAAGACGGGAGT 2102
QY 2227 CACTAGATACTTTTCCAGCCCACTAAGTTATGAACCTGGGGAGCTTTTATCTTTTTCACAT 2286
Db 2103 CACTAGATACTTTTCCAGCCCACTAAGTTATGAACCTGGGGAGCTTTTATCTTTTTCACAT 2162
QY 2287 GCTTTTCTAATTATGCTTGAAGCCCACTACTCTGTTAGGAGAGACATTTCTAGCAAAA 2346
Db 2163 GCTTTTCTAATTATGCTTGAAGCCCACTACTCTGTTAGGAGAGACATTTCTAGCAAAA 2222
QY 2347 GCAGGGCCCATATACCTGAACATAGGAGAGGAACACCCGTTGTTGTTCCCTGCT 2406
Db 2223 GCAGGGCCCATATACCTGAACATAGGAGAGGAACACCCGTTGTTGTTCCCTGCT 2282
QY 2407 TGAGGAAGGAATTAATCTGAAAGTCTGGGCAACAGAGACAAATATGGACGAGCCAAAGA 2466
Db 2283 TGAGGAAGGAATTAATCTGAAAGTCTGGGCAACAGAGACAAATATGGACGAGCCAAAGA 2342
QY 2467 ATGCCCGTCTCTTCAAGTTAAACTAAAGG 2496
Db 2343 ATGCCCGTCTCTTCAAGTTAAACTAAAGG 2372

```

```

RESULT 6
BD267486
LOCUS      2372 bp DNA linear PAT 17-JUL-2003
DEFINITION Endogenous nucleic fragment associated with an autoimmune disease,
            marking method and reagent.
ACCESSION  BD267486
VERSION     BD267486.1 GI:33077254
KEYWORDS   JP 2002534980-A/27.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 2372)
AUTHORS    Baccala,G.P., Mallet,F. and Voisset,C.
TITLE      Endogenous nucleic fragment associated with an autoimmune disease,
            marking method and reagent
JOURNAL     Patent: JP 2002534980-A 27 22-OCT-2002;
            BIO MERIEUX
COMMENT     OS Homo sapiens (human)
            PN JP 2002534980-A/27
            PD 22-OCT-2002
            PE 21-JAN-2000 JP 2000594929
            PR 21-JAN-1999 FR 99/00888
            PI GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
            C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
            G01N33/564,
            PC G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
            Endogenous nucleic fragment associated with an autoimmune CC
            disease,
            CC marking method and reagent
            FH Key Location/Qualifiers
            FT source 1..2372
            FT /organism='Homo sapiens (human)'.
            FT Location/Qualifiers
            1..2372
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            FEATURES
            source

```

/db\_xref="taxon:9606"

ORIGIN

```
Query Match      11.7%; Score 888; DB 2; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGGAGGGAAGAAATCCCTACTGCTTCTGAGAGACTAAGGGAGGCAATTGAGGAAGCG 1566
DB 1383 GGGAGGGAAGAAATCCCTACTGCTTCTGAGAGACTAAGGGAGGCAATTGAGGAAGCG 1442

QY 1567 TGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1626
DB 1443 TGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1502

QY 1627 TCAGTCAGCTGCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTATAGGCCCGGAGCAAAAC 1686
DB 1503 TCAGTCAGCTGCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTATAGGCCCGGAGCAAAAC 1562

QY 1687 TTAGAAACCTTAACTTGAACCTTGGCAACCTTGGTCTTATATAGAGATCAGGAGGAGCAG 1746
DB 1563 TTAGAAACCTTAACTTGAACCTTGGCAACCTTGGTCTTATATAGAGATCAGGAGGAGCAG 1622

QY 1747 GGGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1806
DB 1623 GGGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1682

QY 1807 AGTGACATTTGAGGCTCTGAAAAAGGGAAAAAGCTGGGCAAAATGAAATGCTTAATAGGGC 1866
DB 1683 AGTGACATTTGAGGCTCTGAAAAAGGGAAAAAGCTGGGCAAAATGAAATGCTTAATAGGGC 1742

QY 1867 TTGCTTCCAGTGGGTCTACAGGACACTTTTAAAAAGATTTCTCAAGTATAGAGTAAGCC 1926
DB 1743 TTGCTTCCAGTGGGTCTACAGGACACTTTTAAAAAGATTTCTCAAGTATAGAGTAAGCC 1802

QY 1927 GCCCTCTGCTCACTGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1986
DB 1803 GCCCTCTGCTCACTGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1862

QY 1987 ACAAGGCTTTTGTAGTCAGAGGCCACTAAACAGATGATCCAGCAGGAGCTAGGGGTG 2046
DB 1863 ACAAGGCTTTTGTAGTCAGAGGCCACTAAACAGATGATCCAGCAGGAGCTAGGGGTG 1922

QY 2047 CTGGGGCAAGGCCATCCATGCCATCACCTCTCAGAGCCCTGGGTATGCTTACCAT 2106
DB 1923 CTGGGGCAAGGCCATCCATGCCATCACCTCTCAGAGCCCTGGGTATGCTTACCAT 1982

QY 2107 TGAGGGCCAGGAAGGTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACTCTTCTG 2166
DB 1983 TGAGGGCCAGGAAGGTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACTCTTCTG 2042

QY 2167 TCCCGGACAACTGTCTCTCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGACGGGAGT 2226
DB 2043 TCCCGGACAACTGTCTCTCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGACGGGAGT 2102

QY 2227 CACTAGATATTTTCCGAGCCACTAAGTTATGAAGTGGGAGCTTTATCTTTTCACT 2286
DB 2103 CACTAGATATTTTCCGAGCCACTAAGTTATGAAGTGGGAGCTTTATCTTTTCACT 2162

QY 2287 GCTTTTCTAATATGCTTCAAGGCCCACTACTCTGTGTAGGAGAGACATTTCTAGCAAAA 2346
DB 2163 GCTTTTCTAATATGCTTCAAGGCCCACTACTCTGTGTAGGAGAGACATTTCTAGCAAAA 2222

QY 2347 GCAGGGGCCATTTATACCTGAACATPAGGAAGGAACACCCGTTTGTGTCCTGCT 2406
DB 2223 GCAGGGGCCATTTATACCTGAACATPAGGAAGGAACACCCGTTTGTGTCCTGCT 2282

QY 2407 TGAGGAAGAAATTAATCTGAAGTCTGGGCAACAGAGACAAATATGAGCAGGCCAAGA 2466
DB 2283 TGAGGAAGAAATTAATCTGAAGTCTGGGCAACAGAGACAAATATGAGCAGGCCAAGA 2342

QY 2467 ATGCCCGTCTGCTTCAAGTTAACTAAAGG 2496
DB 1983 ATGCCCGTCTGCTTCAAGTTAACTAAAGG 2042
```

DB 2343 ATGCCCGTCTGCTTCAAGTTAACTAAAGG 2372

```
RESULT 7
LOCUS AX000965
DEFINITION Sequence 10 from Patent WO9902696.
ACCESSION AX000965
VERSION AX000965.1 GI:7241207
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2372)
AUTHORS Beseme,F. and Blond,J.
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
JOURNAL DISEASES OR WITH PREGNANCY DISORDERS
PATENT: WO 9902696-A 10 21-JAN-1999;
BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES
source
Location/Qualifiers
1..2372
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match      11.7%; Score 888; DB 2; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGGAGGGAAGAAATCCCTACTGCTTCTGAGAGACTAAGGGAGGCAATTGAGGAAGCG 1566
DB 1383 GGGAGGGAAGAAATCCCTACTGCTTCTGAGAGACTAAGGGAGGCAATTGAGGAAGCG 1442

QY 1567 TGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1626
DB 1443 TGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1502

QY 1627 TCAGTCAGCTGCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTATAGGCCCGGAGCAAAAC 1686
DB 1503 TCAGTCAGCTGCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTATAGGCCCGGAGCAAAAC 1562

QY 1687 TTAGAAACCTTAACTTGAACCTTGGCAACCTTGGTCTTATATAGAGATCAGGAGGAGCAG 1746
DB 1563 TTAGAAACCTTAACTTGAACCTTGGCAACCTTGGTCTTATATAGAGATCAGGAGGAGCAG 1622

QY 1747 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCCACTGCTTATGTCATGACCCCTCAGGCA 1806
DB 1623 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCCACTGCTTATGTCATGACCCCTCAGGCA 1682

QY 1807 AGTGACATTTTGGAGGCTCTGAAAAAGGGAAAAAGCTGGGCAAAATGAAATGCTTAAAGGCG 1866
DB 1683 AGTGACATTTTGGAGGCTCTGAAAAAGGGAAAAAGCTGGGCAAAATGAAATGCTTAAAGGCG 1742

QY 1867 TTGCTTCCAGTGGGTCTACAGGACACTTTTAAAAAAGATTTGTCCAAGTATAGAGTAAGCC 1926
DB 1743 TTGCTTCCAGTGGGTCTACAGGACACTTTTAAAAAAGATTTGTCCAAGTATAGAGTAAGCC 1802

QY 1927 GCCCTCTGCTCACTGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1986
DB 1803 GCCCTCTGCTCACTGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1862

QY 1987 ACAAGGCTTTTGTAGTCAGAGGCCACTAAACAGATGATCCAGCAGGAGCTAGGGGTG 2046
DB 1863 ACAAGGCTTTTGTAGTCAGAGGCCACTAAACAGATGATCCAGCAGGAGCTAGGGGTG 1922

QY 2047 CTGGGGCAAGGCCATCCATGCCATCACCTCTCAGAGCCCTGGGTATGCTTACCAT 2106
DB 1923 CTGGGGCAAGGCCATCCATGCCATCACCTCTCAGAGCCCTGGGTATGCTTACCAT 1982

QY 2107 TGAGGGCCAGGAAGGTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACTCTTCTG 2166
DB 1983 TGAGGGCCAGGAAGGTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACTCTTCTG 2042

QY 2167 TCCCGGACAACTGTCTCTCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGACGGGAGT 2226
DB 2043 TCCCGGACAACTGTCTCTCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGACGGGAGT 2102

QY 2227 CACTAGATATTTTCCGAGCCACTAAGTTATGAAGTGGGAGCTTTATCTTTTCACT 2286
DB 2103 CACTAGATATTTTCCGAGCCACTAAGTTATGAAGTGGGAGCTTTATCTTTTCACT 2162

QY 2287 GCTTTTCTAATATGCTTCAAGGCCCACTACTCTGTGTAGGAGAGACATTTCTAGCAAAA 2346
DB 2163 GCTTTTCTAATATGCTTCAAGGCCCACTACTCTGTGTAGGAGAGACATTTCTAGCAAAA 2222

QY 2347 GCAGGGGCCATTTATACCTGAACATPAGGAAGGAACACCCGTTTGTGTCCTGCT 2406
DB 2223 GCAGGGGCCATTTATACCTGAACATPAGGAAGGAACACCCGTTTGTGTCCTGCT 2282

QY 2407 TGAGGAAGAAATTAATCTGAAGTCTGGGCAACAGAGACAAATATGAGCAGGCCAAGA 2466
DB 2283 TGAGGAAGAAATTAATCTGAAGTCTGGGCAACAGAGACAAATATGAGCAGGCCAAGA 2342

QY 2467 ATGCCCGTCTGCTTCAAGTTAACTAAAGG 2496
DB 1983 ATGCCCGTCTGCTTCAAGTTAACTAAAGG 2042
```



Qy	2167	TCCCGGACAACTGCTCCAGATCTGCTCACTATTTCTGAGGGGTCCNTAAGACGGGCGT	2226
Db	2043	TCCCGGACAACTGCTCCAGATCTGCTCACTATTTCTGAGGGGTCCNTAAGACGGGCGT	2102
Qy	2227	CACCTAGATACCTTTTCCAGCCACTAAGTTATGAATCTGGGAGGCTTTATTTCTTTTACAT	2286
Db	2103	CACCTAGATACCTTTTCCAGCCACTAAGTTATGAATCTGGGAGGCTTTATTTCTTTTACAT	2162
Qy	2287	GCCTTTCTAATATGCTTGAAGCCCACTACCTTTGTTAGGAGAGACATTTCTAGCAAAA	2346
Db	2163	GCCTTTCTAATATGCTTGAAGCCCACTACCTTTGTTAGGAGAGACATTTCTAGCAAAA	2222
Qy	2347	GCAGGGCCATTATACACCTGACATAGGAGAGGAAACACCGCTTTGTTGTTGTTGCTT	2406
Db	2223	GCAGGGCCATTATACACCTGACATAGGAGAGGAAACACCGCTTTGTTGTTGTTGCTT	2282
Qy	2407	TGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGACAAATATGAGCGAGCAAGA	2466
Db	2283	TGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGACAAATATGAGCGAGCAAGA	2342
Qy	2467	ATGCCGCTCCTGTTCAAGTTAAACTAAAG	2496
Db	2343	ATGCCGCTCCTGTTCAAGTTAAACTAAAG	2372
RESULT 8			
AX027479			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
JOURNAL			
FEATURES			
source			
ORIGIN			

Query Match			
Best Local Similarity			
Matches			
Qy	1507	GGGAGGAGAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCAATGAGGAGCG	1566
Db	1383	GGGAGGAGAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCAATGAGGAGCG	1442
Qy	1567	TGCTCTCTGCTACCTGACTTCTTGAAGCCCACTAATCTTAAAGCGTAAGTTATCAC	1626
Db	1443	TGCTCTCTGCTACCTGACTTCTTGAAGCCCACTAATCTTAAAGCGTAAGTTATCAC	1502
Qy	1627	TCAGTCAGCTGCGAGACATTAAGAAAACTTCAAAAGTCTGCGTAGCGCCGGAGCAAAAC	1686
Db	1503	TCAGTCAGCTGCGAGACATTAAGAAAACTTCAAAAGTCTGCGTAGCGCCGGAGCAAAAC	1562
Qy	1687	TTAGAACCTTATGACTTGGCACTGCTGTTTATTAATAGATCAGGAGGAGCAG	1746
Db	1563	TTAGAACCTTATGACTTGGCACTGCTGTTTATTAATAGATCAGGAGGAGCAG	1622
Qy	1747	CGGGAACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCA	1806
Db	1623	CGGGAACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCA	1682

Qy	1807	AGTGGACTTTGGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATGAAATGCTTAATAGGGC	1866
Db	1683	AGTGGACTTTGGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATGAAATGCTTAATAGGGC	1742
Qy	1867	TTGCTTTCCAGTGGGCTCAAGGACACTTTTAAAAAGATTGTCCAAGTAAAGTAAAGCC	1926
Db	1743	TTGCTTTCCAGTGGGCTCAAGGACACTTTTAAAAAGATTGTCCAAGTAAAGTAAAGCC	1802
Qy	1927	GCCCTTTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG	1986
Db	1803	GCCCTTTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG	1862
Qy	1987	ACAAAGGTCTTTTGAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGT	2046
Db	1863	ACAAAGGTCTTTTGAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGT	1922
Qy	2047	CCTGGGGCAAGGCCATCCATGCCATCACTCCAGAGCCCTGCGTATGCTTGACCAT	2106
Db	1923	CCTGGGGCAAGGCCATCCATGCCATCACTCCAGAGCCCTGCGTATGCTTGACCAT	1982
Qy	2107	TGAGGCCCAAGGAGTGTCTCTGGAACACTGCTGCGGTCTTCTTACTTACTTCTG	2166
Db	1983	TGAGGCCCAAGGAGTGTCTCTGGAACACTGCTGCGGTCTTCTTACTTACTTCTG	2042
Qy	2167	TCCCGGACAACTGCTCCTCAGATCTGCTCACTATTTCTGAGGGGTCCNTAAGACGGGCGT	2226
Db	2043	TCCCGGACAACTGCTCCTCAGATCTGCTCACTATTTCTGAGGGGTCCNTAAGACGGGCGT	2102
Qy	2227	CACCTAGATCTTTTCCAGCCACTAAGTTATGAATCTGAGGAGCTTTATTTCTTTTACAT	2286
Db	2103	CACCTAGATCTTTTCCAGCCACTAAGTTATGAATCTGAGGAGCTTTATTTCTTTTACAT	2162
Qy	2287	GCTTTTCTAATATGCTTGAAGCCCACTACCTTCTTAGGGAGAGACATTTCTAGCAAAA	2346
Db	2163	GCTTTTCTAATATGCTTGAAGCCCACTACCTTCTTAGGGAGAGACATTTCTAGCAAAA	2222
Qy	2347	GCAGGGCCATTATACCTGAAACATAGAGAGAAACACCGCTTTGTTGTTGTTGCTGCT	2406
Db	2223	GCAGGGCCATTATACCTGAAACATAGAGAGAAACACCGCTTTGTTGTTGTTGCTGCT	2282
Qy	2407	TGAGGAAGGAATTAATCTGGAAGTCTGGGCAACAGAGGACAAATATGAGCGAGCAAGA	2466
Db	2283	TGAGGAAGGAATTAATCTGGAAGTCTGGGCAACAGAGGACAAATATGAGCGAGCAAGA	2342
Qy	2467	ATGCCGCTCCTGTTCAAGTTAAACTAAAG	2496
Db	2343	ATGCCGCTCCTGTTCAAGTTAAACTAAAG	2372

RESULT 9	AF072504	2372 bp	mRNA	linear	PRI 10-FEB-1999
LOCUS	AF072504				
DEFINITION	Homo sapiens endogenous retrovirus W sequence.				
ACCESSION	AF072504				
VERSION	AF072504.1	GI:4262288			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2372)				
AUTHORS	Blond, J.-L., Beseme, F., Duret, L., Bouton, O., Bedin, F., Perron, H.,				
	Mandrand, B., and Mallet, F.				
TITLE	Molecular characterization and placental expression of HERV-W, a				
JOURNAL	new human endogenous retrovirus family				
PUBMED	J. Virol. 73 (2), 1175-1185 (1999)				
REFERENCE	2 (bases 1 to 2372)				
AUTHORS	Blond, J.-L., Beseme, F. and Mallet, F.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-JUN-1998) UMI03 CNRS-bioMerieux, ENS Lyon, 46 allée				
	d'Italie, Lyon, Cedex 07 69364, France				

```

FEATURES
source
1..2372
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl.44.4"
/tissue_type="placenta"
/clone_lib="Clontech 5'-stretch plus library, Cat number
HU5014a"
misc_feature
1..2372
/note="corresponds to gag and pol"
repeat_region
1..2372
/rpt_family="HERV-W"
5'UTR
<1..6
/note="R"
7..460
/note="US"
464..481
/note="putative primer binding site for tRNA-W"
ORIGIN
Query Match 11.7%; Score 888; DB 5; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1507 GGGAGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCAATGAGGAGCG 1566
DB 1383 GGGAGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCAATGAGGAGCG 1442
QY 1567 TGCCTCTCTGTCACCTGACTCTTCTGAAGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1626
DB 1443 TGCCTCTCTGTCACCTGACTCTTCTGAAGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1502
QY 1627 TCAGTCAGCTGCAGACATAGAAAAAAGCTTCAAAAGCTCTGCCGTAGGCCCGGAGCAAAAC 1686
DB 1503 TCAGTCAGCTGCAGACATAGAAAAAAGCTTCAAAAGCTCTGCCGTAGGCCCGGAGCAAAAC 1562
QY 1687 TTAGAAACCCATTGAAGCTTGGCACTGCGTTTATTAATAGATCAGAGGAGCAG 1746
DB 1563 TTAGAAACCCATTGAAGCTTGGCACTGCGTTTATTAATAGATCAGAGGAGCAG 1622
QY 1747 GCGGAACAGACAAACGGGATTAAAAAAGGCCACCGCTTTAGTCATCACCTCAGGCA 1806
DB 1623 GCGGAACAGACAAACGGGATTAAAAAAGGCCACCGCTTTAGTCATCACCTCAGGCA 1682
QY 1807 AGTGGACTTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTGAATCCCTAATAGGCG 1866
DB 1683 AGTGGACTTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTGAATCCCTAATAGGCG 1742
QY 1867 TTGCTTCCAGTGGCGTCTCAGGACACATTTAAAAAAGATTGTCGAAGTAGAAGTAAGCC 1926
DB 1743 TTGCTTCCAGTGGCGTCTCAGGACACATTTAAAAAAGATTGTCGAAGTAGAAGTAAGCC 1802
QY 1927 GCCCTTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACTGCCCCAGGGG 1986
DB 1803 GCCCTTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACTGCCCCAGGGG 1862
QY 1987 ACAAAAGTCTTTTGAAGTCAGAACCCACTAAACAGATGATCCAGCAGCAGGACTGAGGGTG 2046
DB 1863 ACAAAAGTCTTTTGAAGTCAGAACCCACTAAACAGATGATCCAGCAGCAGGACTGAGGGTG 1922
QY 2047 CTGGGGCAAGCCCATCCATGCCATCACTCCATCAGAGCCCTGGGTATGCTTGACCAT 2106
DB 1923 CTGGGGCAAGCCCATCCATGCCATCACTCCATCAGAGCCCTGGGTATGCTTGACCAT 1982
QY 2107 TGAGGGCCAGGAAGTTGTCTCTGACACCTGTCGGGTCTTCTTAGTCTTACTCTTCG 2166
DB 1983 TGAGGGCCAGGAAGTTGTCTCTGACACCTGTCGGGTCTTCTTAGTCTTACTCTTCG 2042
QY 2167 TCCCGGACAACTGCTCCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGGCAGT 2226
DB 2043 TCCCGGACAACTGCTCCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGGCAGT 2102

```

```

QY 2227 CACTAGACTATTTTCCAGCCACTAAGCTTATGAATCGGAGCTTTATTTTTCACAT 2286
DB 2103 CACTAGACTATTTTCCAGCCACTAAGCTTATGAATCGGAGCTTTATTTTTCACAT 2162
QY 2287 GCTTTTCTAATTAATGCTTTGAAAGCCCACTACCTTTGTTAGGAGAGACATTTAGCAAAA 2346
DB 2163 GCTTTTCTAATTAATGCTTTGAAAGCCCACTACCTTTGTTAGGAGAGACATTTAGCAAAA 2222
QY 2347 GCAGGGGCATTAATACCTGAACATAGGAGAGAAACACCCGTTGTTGTTNCCCCTGCT 2406
DB 2223 GCAGGGGCATTAATACCTGAACATAGGAGAGAAACACCCGTTGTTGTTNCCCCTGCT 2282
QY 2407 TGAGGAAGGAATTAATCTGAAGCTCTGGCAACAGAGGACAAATATGACGAGCCAAAGA 2466
DB 2283 TGAGGAAGGAATTAATCTGAAGCTCTGGCAACAGAGGACAAATATGACGAGCCAAAGA 2342
QY 2467 ATGCCCGTCTGTTTCAAGTTAAACTAAAGG 2496
DB 2343 ATGCCCGTCTGTTTCAAGTTAAACTAAAGG 2372
RESULT 10
BD221806
LOCUS
DEFINITION
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses.
ACCESSION
BD221806
VERSION
BD221806.1 GI:33031576
KEYWORDS
JP 2002518051-A/1.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 2599)
AUTHORS
Alliel,P.M., Perin,J.P. and Rieger,F.
TITILE
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL
PATENT: JP 2002518051-A 1 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT
OS Homo sapiens (human)
PN JP 2002518051-A/1
PD 25-JUN-2002
PF 23-JUN-1999 JP 2000556036
PR 23-JUN-1998 FR 98/07920
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
A61P25/00.
PC A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC Nucleic
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1..2599
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..2599
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 10.3%; Score 779; DB 2; Length 2599;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;
QY 4802 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAGGCACTAGTTTCATGAA 4861
DB 151 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAGGCACTAGTTTCATGAA 210
QY 4862 TAAATCCCAAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 4921
DB 211 TAAATCCCAAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 270

```



Query Match		10.3%; Score 779; DB 2; Length 2599;	
Best Local Similarity		98.9%; Pred. No. 0;	
Matches 1899; Conservative		0; Mismatches 20; Indels 1; Gaps 1;	
QY	4802	AGGCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTATAAAGGCATAGTTCAATGAAA	4861
DB	151	AGGCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTATAAAGGCATAGTTCAATGAAA	210
QY	4862	TAATTCAGGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG	4921
DB	211	TAATTCAGGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG	270
QY	4922	CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATCGCGCT	4981
DB	271	CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATCGCGCT	330
QY	4982	GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA	5041
DB	331	GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA	390
QY	5042	AAAAGCAAAACCCAGGAAACCCACTCACATGGCCTGTCTGTGGCTATAGCCCTTAAAAA	5101
DB	391	AAAAGCAAAACCCAGGAAACCCACTCACATGGCCTGTCTGTGGCTATAGCCCTTAAAAA	450
QY	5102	GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAGGC	5161
DB	451	GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAGGC	510
QY	5162	CTTTCATAACCAATGACCTTGTGCTTACCCAGGACAGCAACTTAGTTGACAGATCAC	5221
DB	511	CTTTCATAACCAATGACCTTGTGCTTACCCAGGACAGCAACTTAGTTGACAGATCAC	570
QY	5222	CTCCTTAGCAATATCAACAAGTCTTAAACATTAACAAGAACCTATCCCTGAGAGA	5281
DB	571	CTCCTTAGCAATATCAACAAGTCTTAAACATTAACAAGAACCTATCCCTGAGAGA	630
QY	5282	GGGAAAGAACTATTCCACCWGTGACATGTTAGTCAAGTCCCTTCYCTTAATTC	5341
DB	631	GGGAAAGAACTATTCCACCWGTGACATGTTAGTCAAGTCCCTTCYCTTAATTC	690
QY	5342	CCCATCCCTAGATACATCTGGGAAGGACCCCTACCAGTCACTTTATYATACCCCAACTGC	5401
DB	691	CCCATCCCTAGATACATCTGGGAAGGACCCCTACCAGTCACTTTATYATACCCCAACTGC	750
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTTGAGTCGCAAAATCTGTGATCT	5461
DB	751	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTTGAGTCGCAAAATCTGTGATCT	810
QY	5462	GCCAAAGGAACCTGAAAATCCAGGAGACACCGCTAGCTATTCTGTGAACCTTAGAGGA	5521
DB	811	GCCAAAGGAACCTGAAAATCCAGGAGACACCGCTAGCTATTCTGTGAACCTTAGAGGA	870
QY	5522	TTTGGCCCTGCTCTTCAAAACAACACGAGGAAAGTAACATAAATCAATAATCCCCA	5581
DB	871	TTTGGCCCTGCTCTTCAAAACAACACGAGGAAAGTAACATAAATCAATAATCCCCA	929
QY	5582	TGSCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCCTSTTTTCACTCTCACTG	5641
DB	930	TGSCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCCTSTTTTCACTCTCACTG	989
QY	5642	CACCCCTCATGCGCTGTATGACCAGTAGTCTCCCTTACCMAGAGTTTCTATGAGAA	5701
DB	990	CACCCCTCATGCGCTGTATGACCAGTAGTCTCCCTTACCMAGAGTTTCTATGAGAA	1049
QY	5702	TGCAAGGCTCCGGAATATTGATGCCCATCGTATAGGAGTCTTCTTAAGGGAACCCCA	5761
DB	1050	TGCAAGGCTCCGGAATATTGATGCCCATCGTATAGGAGTCTTCTTAAGGGAACCCCA	1109
QY	5762	CTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTGTGCCACTCTTTGATGC	5821
DB	1110	CTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTGTGCCACTCTTTGATGC	1169
QY	5822	ATGCAAAATACATCATTTATTGGACAGGAAAATGATTAATCCTAGTTGCTCTGGAGGACTTG	5881
DB	1170	ATGCAAAATACATCATTTATTGGACAGGAAAATGATTAATCCTAGTTGCTCTGGAGGACTTG	1229
QY	5882	GAGTCACTGTCTGTGGACTTACTTACCCAAAACCTGGTATGTCTGATGGGGTGGAGTTC	5941
DB	1230	GAGTCACTGTCTGTGGACTTACTTACCCAAAACCTGGTATGTCTGATGGGGTGGAGTTC	1289
QY	5942	AAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCAACTACCCSGGGTACATG	6001
DB	1290	AAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCAACTACCCSGGGTACATG	1349
QY	6002	GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAACCTAGTAAACCCCTCCGTACCC	6061
DB	1350	GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAACCTAGTAAACCCCTCCGTACCC	1409
QY	6062	ATATCGCCTCGTAAAGCCTATTAAATACCAACCTCACTGGGCTCATGAGGTCTCGGCC	6121
DB	1410	ATATCGCCTCGTAAAGCCTATTAAATACCAACCTCACTGGGCTCATGAGGTCTCGGCC	1469
QY	6122	AAAACCTTACTAACTGTGGATATGCTCCCTGAACTTCARGCCATATGTTCAATCC	6181
DB	1470	AAAACCTTACTAACTGTGGATATGCTCCCTGAACTTCARGCCATATGTTCAATCC	1529
QY	6182	CTGTACCTGAAACAATGGAAACAACCTTCAGCAGAGAAATAAACAACCACTTCCGTTTAGTAG	6241
DB	1530	CTGTACCTGAAACAATGGAAACAACCTTCAGCAGAGAAATAAACAACCACTTCCGTTTAGTAG	1589
QY	6242	GACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTTA	6301
DB	1590	GACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTTA	1649
QY	6302	GCAATACCTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTTCTCCCAACAAA	6361
DB	1650	GCAATACCTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTTCTCCCAACAAA	1709
QY	6362	TAGTCTGCTACCTCAGGAATATTTTGTGTGTGGTACCTCAGCCCTATCGTTGTTGA	6421
DB	1710	TAGTCTGCTACCTCAGGAATATTTTGTGTGTGGTACCTCAGCCCTATCGTTGTTGA	1769
QY	6422	ATGGCTCTTCAAGATCTAATGTGCTCTCATTTAGTGCCCTCCCTATGRCATCTACA	6481
DB	1770	ATGGCTCTTCAAGATCTAATGTGCTCTCATTTAGTGCCCTCCCTATGRCATCTACA	1829
QY	6482	CTGAACAGATTTATACAGTTATGTCTATCTTAGCCCCGCAACAAAGAGTACCCATTC	6541
DB	1830	CTGAACAGATTTATACAGTTATGTCTATCTTAGCCCCGCAACAAAGAGTACCCATTC	1889
QY	6542	TTCCCTTTTGTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA	6601
DB	1890	TTCCCTTTTGTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA	1949
QY	6602	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG	6661
DB	1950	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG	2009
QY	6662	TCGCCGACTCCCTGCTGCTTGAAGATCACTTAATCTCCCTAGCAGCAGTAGTCTCTTC	6721
DB	2010	TCGCCGACTCCCTGCTGCTTGAAGATCACTTAATCTCCCTAGCAGCAGTAGTCTCTTC	2069
RESULT 12			
AX007978			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1			
AX007978			
Sequence 1 from Patent WO9967395.			
AX007978.1			
GI:9995675			
linear			
PAT 06-SEP-2000			

AUTHORS TITLE JOURNAL FEATURES source	Perin,J.P., Rieger,F. and Alliel,P.M. Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses Patent: WO 9967395-A 1 29-DEC-1999; INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER FRANCOIS Location/Qualifiers		1. .2599 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	
	ORIGIN			
	Query Match 10.3%; Score 779; DB 2; Length 2599; Best Local Similarity 98.9%; Pred. No. 0; Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;			
QY	4802	AGSCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA	4861	
DB	151	AGSCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA	210	
QY	4862	TAATTCGAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG	4921	
DB	211	TAATTCGAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG	270	
QY	4922	CCACAGTACCCAGGAGGATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT	4981	
DB	271	CCACAGTACCCAGGAGGATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT	330	
QY	4982	GAAGGCCACAGTCTCAGGGAAGGTGAGAAAATGAATGAAAVACTCAAGAGGACATCTAA	5041	
DB	331	GAAGGCCACAGTCTCAGGGAAGGTGAGAAAATGAATGAAAVACTCAAGAGGACATCTAA	390	
QY	5042	AAAAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTGGCTTATAGCCTTAAAAA	5101	
DB	391	AAAAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTGGCTTATAGCCTTAAAAA	450	
QY	5102	GAATCTGCACTTTCCTCCAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	5161	
DB	451	GAATCTGCACTTTCCTCCAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	510	
QY	5162	CTTTCAATACCAATGACTTGTCTTGACCCAGAGACGCACTTAGTTGCGACATCAC	5221	
DB	511	CTTTCAATACCAATGACTTGTCTTGACCCAGAGACGCACTTAGTTGCGACATCAC	570	
QY	5222	CTCTTAGCCAAATATCAACAGTTCTTAAAAATTTACAGGAACCTATCCCTTGAGAGA	5281	
DB	571	CTCTTAGCCAAATATCAACAGTTCTTAAAAATTTACAGGAACCTATCCCTTGAGAGA	630	
QY	5282	GGGAAAGAACTATTCACCCWGTGACATGGPATTTAGTCAAGTCCCTTCYCTCTAATTTC	5341	
DB	631	GGGAAAGAACTATTCACCCWGTGACATGGPATTTAGTCAAGTCCCTTCYCTCTAATTTC	690	
QY	5342	CCCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCAATTTATYATCCCACTGC	5401	
DB	691	CCCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCAATTTATCTACCCCACTGC	750	
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATTTAGTCAAAATCTCTGGATCT	5461	
DB	751	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATTTAGTCAAAATCTCTGGATCT	810	
QY	5462	GCCAAAGGAACTTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGGA	5521	
DB	811	GCCAAAGGAACTTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGGA	870	
QY	5522	TTTGCCCTGCTCTTCAAAACAAACAGGAGGAAAGTAACATAAATCAATAATCCCCCA	5581	
DB	871	TTTGCCCTGCTCTTCAAAACAAACAGGAGGAAAGTAACATAAATCAATAATCCCCCA	929	
QY	5582	TGGSCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCCTSTTTTCACTCTCACTG	5641	
DB	930	TGGSCTCCCTTATCATATTTTCTCTTTACTGTTCTTTTACCCTCTTTTCACTCTCACTG	989	
QY	5642	CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAA	5701	
DB	990	CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAA	1049	
QY	5702	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAGGGAACCCCA	5761	
DB	1050	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCA	1109	
QY	5762	CTTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGC	5821	
DB	1110	CTTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGC	1169	
QY	5822	ATGCAATATCTCATTTATGACAGGAAATGTAATCTTAGTTGCTCGAGGACTTG	5881	
DB	1170	ATGCAATATCTCATTTATGACAGGAAATGTAATCTTAGTTGCTCGAGGACTTG	1229	
QY	5882	GAGTCACTGCTGTGGACTTACTTCAACCAACTGTTATGCTGATGGGGTGGAGTTC	5941	
DB	1230	GAGTCACTGCTGTGGACTTACTTCAACCAACTGTTATGCTGATGGGGTGGAGTTC	1289	
QY	5942	AAGATCAGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACATG	6001	
DB	1290	AAGATCAGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACATG	1349	
QY	6002	GCACCTCTAGCCCTACAAAGAGTAGATCTCTCAAACTACATGAAACCTCTCCGTACCC	6061	
DB	1350	GCACCTCTAGCCCTACAAAGAGTAGATCTCTCAAACTACATGAAACCTCTCCGTACCC	1409	
QY	6062	ATATCGCTGCTGTAAGCTTATTTAATACCACCTCACTGGGCTCCATGAGGTCTCGGCC	6121	
DB	1410	ATATCGCTGCTGTAAGCTTATTTAATACCACCTCACTGGGCTCCATGAGGTCTCGGCC	1469	
QY	6122	AAAACTCTAACTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC	6181	
DB	1470	AAAACTCTAACTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC	1529	
QY	6182	CTGTACTGAAACATAGGAAACAACTTCAGCAGCAAGAAATAAACACCACTTCGGTTTAGTAG	6241	
DB	1530	CTGTACTGAAACATAGGAAACAACTTCAGCAGCAAGAAATAAACACCACTTCGGTTTAGTAG	1589	
QY	6242	GACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCAGCTGTGTAATAATTTA	6301	
DB	1590	GACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCAGCTGTGTAATAATTTA	1649	
QY	6302	GCAATACTACATACACCAACTCCCAATGATCAGTGGGTAACTCTCCACACAAA	6361	
DB	1650	GCAATACTACATACACCAACTCCCAATGATCAGTGGGTAACTCTCCACACAAA	1709	
QY	6362	TAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA	6421	
DB	1710	TAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA	1769	
QY	6422	ATGGCTCTCAGAACTATGTCCTCTCAATCTTAGTGCCCTCCATGRCATCTACA	6481	
DB	1770	ATGGCTCTCAGAACTATGTCCTCTCAATCTTAGTGCCCTCCATGRCATCTACA	1829	
QY	6482	CTGAAACAGATTATACAGTTATGTATCTAAGCCCGCAACAAAAGAGTACCCATTTC	6541	
DB	1830	CTGAAACAGATTATACAGTTATGTATCTAAGCCCGCAACAAAAGAGTACCCATTTC	1889	
QY	6542	TTCTTTTGTATTAGGAGCAGAGTGTAGTGCACTAGGTACTGGCATTTGCGGTATCA	6601	
DB	1890	TTCTTTTGTATTAGGAGCAGAGTGTAGTGCACTAGGTACTGGCATTTGCGGTATCA	1949	
QY	6602	GAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGAGCATGGACGGG	6661	
DB	1950	GAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGAGCATGGACGGG	2009	
QY	6662	TGCGCGACTCCCTGCTGCTTGAAGATCACTTAACTCCCTAGCAGAGTAGTCTCTTC	6721	
DB	2010	TGCGCGACTCCCTGCTGCTTGAAGATCACTTAACTCCCTAGCAGAGTAGTCTCTTC	2069	

RESULT 13	AR612522	Sequence 21 from patent US 6828097.	8523 bp	DNA	linear	PAT 15-DEC-2004
LOCUS	AR612522	Sequence 21 from patent US 6828097.	8523 bp	DNA	linear	PAT 15-DEC-2004
DEFINITION	AR612522	Sequence 21 from patent US 6828097.	8523 bp	DNA	linear	PAT 15-DEC-2004
ACCESSION	AR612522	Sequence 21 from patent US 6828097.	8523 bp	DNA	linear	PAT 15-DEC-2004
VERSION	AR612522.1	GI:56668344	8523 bp	DNA	linear	PAT 15-DEC-2004
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 8523)					
AUTHORS	Knoll,J.H.M. and Rogan,P.K.					
TITLE	Single copy genomic hybridization probes and method of generating same					
JOURNAL	Patent: US 6828097-A 21 07-DEC-2004;					
FEATURES	The Childrens Mercy Hospital; Kansas City, MO					
source	Location/Qualifiers					
	1. 8523					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN						
	Query Match	10.3%; Score 779; DB 2; Length 8523;				
	Best Local Similarity	98.9%; Pred. No. 0;				
	Matches 1899; Conservative	0; Mismatches 20; Indels 1; Gaps 1;				
Qy	4802	AGGCCTTCCCTGTGAGCACAGAAAGCCCAAGAGGTAAATAAAGGCACATAGTTTCATGAAA	4861			
Db	6071	AGGCCTTCCCTGTGAGCACAGAAAGCCCAAGAGGTAAATAAAGGCACATAGTTTCATGAAA	6130			
Qy	4862	TAATTTCCAGATTCCGACTTCCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGG	4921			
Db	6131	TAATTTCCAGATTCCGACTTCCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGG	6190			
Qy	4922	CCACAGTAACCCAGGGAGTATCCACAGCGTTAGGTATACGATATCACTTACACTGCGCCT	4981			
Db	6191	CCACAGTAACCCAGGGAGTATCCACAGCGTTAGGTATACGATATCACTTACACTGCGCCT	6250			
Qy	4982	GAAGGCCACAGTCTCTCAGGGAAGGTTCAGAAAATGAATGAAAVACTTCAAGGACATCTAA	5041			
Db	6251	GAAGGCCACAGTCTCTCAGGGAAGGTTCAGAAAATGAATGAAAVACTTCAAGGACATCTAA	6310			
Qy	5042	AAAAGCAAAACCCAGGAAACCCACTCATATGGCTCTGTCTGTGGCTTATAGCCCTTAAAAA	5101			
Db	6311	AAAAGCAAAACCCAGGAAACCCACTCATATGGCTCTGTCTGTGGCTTATAGCCCTTAAAAA	6370			
Qy	5102	GAATCTGCAACTTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	5161			
Db	6371	GAATCTGCAACTTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	6430			
Qy	5162	CCTTTCATAACCAATGACTTGTGCTTGACCCAGACAGCCAACTTAGTTGCGACATCAC	5221			
Db	6431	CCTTTCATAACCAATGACTTGTGCTTGACCCAGACAGCCAACTTAGTTGCGACATCAC	6490			
Qy	5222	CTCCTTAGCCAAATATCAACAAGTTCTTAAAAAATTACAAGGAACCTTATCCCTGGAAGA	5281			
Db	6491	CTCCTTAGCCAAATATCAACAAGTTCTTAAAAAATTACAAGGAACCTTATCCCTGGAAGA	6550			
Qy	5282	GGGAAAAAGAACTATTTCACCCWMTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	5341			
Db	6551	GGGAAAAAGAACTATTTCACCCWMTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	6610			
Qy	5342	CCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCGAGTCAATTTATYATACCCCACTGC	5401			
Db	6611	CCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCGAGTCAATTTATYATACCCCACTGC	6670			
Qy	5402	GGTTTAAAGTGGCTCGAGTGGAGTCTTCGATACATCACACTTGGTCAAAATCCCTGGATACT	5461			
Db	6671	GGTTTAAAGTGGCTCGAGTGGAGTCTTCGATACATCACACTTGGTCAAAATCCCTGGATACT	6730			
Qy	5462	GCCAAAGGAACTCTGAAAAATCCAGGAGACACGCTAGCTATTTCCTGTGAACTCTTAGAGGA	5521			
Db	6731	GCCAAAGGAACTCTGAAAAATCCAGGAGACACGCTAGCTATTTCCTGTGAACTCTTAGAGGA	6790			

```

QY 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661
|
|
|
Db 7870 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 7929
|
|
|
QY 6662 TCCCGACTCCCTGGTCACTTCCAGAGATCAAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 6721
|
|
|
Db 7930 TCCCGACTCCCTGGTCACTTCCAGAGATCAAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 7989
|
|
|
RESULT 14
AY101582
LOCUS
DEFINITION
Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVW1 locus,
allele A, complete sequence.
ACCESSION
AY101582
VERSION
AY101582.1 GI:37544399
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 (bases 1 to 10222)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
REFERENCE
2 (bases 1 to 10222)
Mallet,F., Bouton,O. and Oriol,G.
Direct Submission
Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
d'Italie, Lyon 69364 cedex 07, France
LOCATION/Qualifiers
1..10222
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="21"
/isolation_source="PBMC of individual 21"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC"
/note="isolated by PCR; ERVW1 locus, allele A
endogenous_virus: HERV-W"
1..780
Join(248..811,2887..9813)
/note="putative mRNA transcript 1; gag-pol pseudogene"
Join(248..811,7585..9813)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
Join(248..811,9248..9813)
/note="putative mRNA transcript 3; no coding region"
7812..9428
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68161.1"
/db_xref="GI:37544400"
/translation="MALPYHIFLFTLLPSFTLAPPCCRCMTSSSPQYELRWORP
GNIDAPSVRSLSKGTPTFTAHMPNRCVHSATLCHMANTHYMTCKMINPSCGGLGV
TVCWTYFTGTGMSGGVQDOAREKHVKEVISQLTRVGTSSPYGLDLKLHETLRT
HRLVSLFNLTILGLHVSQAQNTMCLPLNFRPVSIPVPEQWNPFSEINTISV
LVGLVSLNLEIHTLSNLKVCFSNTTYTNSQCIRVTPPTQIVCLPSGIFPVCGTSA
YRLNGSSSEMSFLFVPPMTIYTEQDLYSVISKPRKRPVILPFVIGAVLGALG
TGIGITSTQPYKLSQELNGDMERVADSLVTLQDLSLAADVLRNRRALLLTAE
RGGTCLFLGECCYVYVNSGIVTEKVKELRDRIRAREBELRNTGPGLLSOWMPLP
FLGPLAAIILLFLFGPCIFNLLVNFVSSRIENAVKLQEPKMQSKTKIYRRPLDRPAS
RSDVNDIKGTPEISAAQPLLRPNAGSS"
9487..10222

```

LTR

## ORIGIN

```

Query Match 10.3%; Score 779; DB 5; Length 10222;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGTAATAAAGGCACACTAGTTCATGAAA 4861
|
|
|
Db 7034 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGTAATAAAGGCACACTAGTTCATGAAA 7093
|
|
|
QY 4862 TAATTCACAGATTCGGACATTCCTCCAGAGCTTACAGAGTGCACATAGCCCTGCTTCCAGG 4921
|
|
|
Db 7094 TAATTCACAGATTCGGACATTCCTCCAGAGCTTACAGAGTGCACATAGCCCTGCTTCCAGG 7153
|
|
|
QY 4922 CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 4981
|
|
|
Db 7154 CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 7213
|
|
|
QY 4982 GAAGGCCACAGTCTCAGGGAAGGTGCAGAAAATGAATGAAYACTCAAAGGCATCTTAA 5041
|
|
|
Db 7214 GAAGGCCACAGTCTCAGGGAAGGTGCAGAAAATGAATGAAYACTCAAAGGCATCTTAA 7273
|
|
|
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGCTTATAGCCTTAATAA 5101
|
|
|
Db 7274 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGCTTATAGCCTTAATAA 7333
|
|
|
QY 5102 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
|
|
|
Db 7334 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 7393
|
|
|
QY 5162 CTTTCAATAACCAATGACCTTGTGTTGACCCAGACAGCCAACTTAGTTAGTCAGACATCAC 5221
|
|
|
Db 7394 CTTTCAATAACCAATGACCTTGTGTTGACCCAGACAGCCAACTTAGTTAGTCAGACATCAC 7453
|
|
|
QY 5222 CTCCTTAGCCCAATATCAACAGTCTTAAACATTTACAGGAACCTATCCCTGAGAAGA 5281
|
|
|
Db 7454 CTCCTTAGCCCAATATCAACAGTCTTAAACATTTACAGGAACCTATCCCTGAGAAGA 7513
|
|
|
QY 5282 GGGAAAAGAACTATTCCACCCWMTGACATGCTATTAGTCAAGTCCCTTCYCTCTAATTC 5341
|
|
|
Db 7514 GGGAAAAGAACTATTCCACCCWMTGACATGCTATTAGTCAAGTCCCTTCYCTCTAATTC 7573
|
|
|
QY 5342 CCATCCCTAGATACATCTCTGGAAAGGACCTTACCAGTCATTTTATYACCCCACTGC 5401
|
|
|
Db 7574 CCATCCCTAGATACATCTCTGGAAAGGACCTTACCAGTCATTTTATYACCCCACTGC 7633
|
|
|
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACACTTACGTCAAATCTCTGATCT 5461
|
|
|
Db 7634 GGTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACACTTACGTCAAATCTCTGATCT 7693
|
|
|
QY 5462 GCCAAGGAACCTGAAAATCCAGGAGACAAACGCTAGTATTCTGTGAACCTCTAGAGGA 5521
|
|
|
Db 7694 GCCAAGGAACCTGAAAATCCAGGAGACAAACGCTAGTATTCTGTGAACCTCTAGAGGA 7753
|
|
|
QY 5522 TTTGCGCTGCTCTTCAAAACAAACACAGGAGAAAGTAATAAATCATATAATCCCCCA 5581
|
|
|
Db 7754 TTTGCGCTGCTCTTCAAAACAAACACAGGAGAAAGTAATAAATCATATAATCCCCCA 7812
|
|
|
QY 5582 TGGSCCTCCCTTATCATATTTTCTTATSTGTTTATYACCTTCTTCTCTCTCTCTCT 5641
|
|
|
Db 7813 TGGSCCTCCCTTATCATATTTTCTTATSTGTTTATYACCTTCTTCTCTCTCTCTCTCT 7872
|
|
|
QY 5642 CACCCCTCCATGCGCGCTGATGACAGTACCTCCCTYACCMAGAGTTCTTATGAGAA 5701
|
|
|
Db 7873 CACCCCTCCATGCGCGCTGATGACAGTACCTCCCTYACCMAGAGTTCTTATGAGAA 7932
|
|
|
QY 5702 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761
|
|
|
Db 7932 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 7992
|
|
|
QY 5762 CCTTCACTGCCCCACACCCATATGCCCGCAACTGTCTATCACTCTGCCACTCTTTGATGC 5821
|
|
|
Db 7993 CCTTCACTGCCCCACACCCATATGCCCGCAACTGTCTATCACTCTGCCACTCTTTGATGC 8052
|
|
|

```







Db 7214 GAAGGCCACAGTCTCTAGGAGAGGTTCAGAAATGAATGAACACTCAAAGGACATCTAA 7273  
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCATAGGCTGTCTGTGCTTATAGCCTTAAAAA 5101  
Db 7274 AAAAGCAAAACCCAGGAAACCCACCTCATAGGCTGTCTGTGCTTATAGCCTTAAAAA 7333  
QY 5102 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGCAAACTGTATGGAGGC 5161  
Db 7334 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGCAAACTGTATGGAGGC 7393  
QY 5162 CCTTCAATCAATAGCTTTGTGCTTGACCCCAAGACAGCAACTTAGTTGCGACATCAC 5221  
Db 7394 CCTTCAATCAATAGCTTTGTGCTTGACCCCAAGACAGCAACTTAGTTGCGACATCAC 7453  
QY 5222 CTCCTTAGCCAAATATCAAGAAGTTCTTAAAAATTAACAAGAACTTATCCCTGAGAAGA 5281  
Db 7454 CTCCTTAGCCAAATATCAAGAAGTTCTTAAAAATTAACAAGAACTTATCCCTGAGAAGA 7513  
QY 5282 GGGAAAGAACTATTCACCCCMWGTGACATGCTATAGTCAAGTCCCTTCYCTTAATTC 5341  
Db 7514 GGGAAAGAACTATTCACCCCTTGTGACATGGTATAGTCAAGTCCCTTCCTTAATTC 7573  
QY 5342 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCOCAGTCAATTTATATACCCCACTGC 5401  
Db 7574 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCOCAGTCAATTTATCTACCCCACTGC 7633  
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGAATACATCACTTTGAGTCAAACTCTGGATAC 5461  
Db 7634 GGTAAAGTGGCTGGAGTGGAGTCTTGGAATACATCACTTTGAGTCAAACTCTGGATAC 7693  
QY 5462 GCCAAGGAACCTGAAAATCCAGAGAACACGCTAGCTATCTCTGTGAACCTCTAGAGGA 5521  
Db 7694 GCCAAGGAACCTGAAAATCCAGAGAACACGCTAGCTATCTCTGTGAACCTCTAGAGGA 7753  
QY 5522 TTTGCGCTCTCTTCAAAACAAACAGAGAGAACTAACTAAATCAATATCCCA 5581  
Db 7754 TTTGCGCTCTCTTCAAAACAAACAGAGAGAACTAACTAAATCAATATCCCA 7812  
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCCTSTTTTCACTCTCACTG 5641  
Db 7813 TGGSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCCTCTTTTCACTCTCACTG 7872  
QY 5642 CACCCCTCATGCGCTGTATGACAGTAGTCTCCCTTACCMAGAGTTTCTATGAGAA 5701  
Db 7873 CACCCCTCATGCGCTGTATGACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGAA 7932  
QY 5702 TGCAGAGTCCCGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAGGGAACCCCA 5761  
Db 7933 TGCAGAGTCCCGAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCA 7992  
QY 5762 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 5821  
Db 7993 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 8052  
QY 5822 ATGCAATACTCTATTTGACAGGAAAATGATTAATCTAGTTGCTCTGGAGGACTTG 5881  
Db 8053 ATGCAATACTCTATTTGACAGGAAAATGATTAATCTAGTTGCTCTGGAGGACTTG 8112  
QY 5882 GAGTCACTGCTCTTGGACTTACTTACCCAACTGCTATGCTCTGATGGGGTGGAGTTC 5941  
Db 8113 GAGTCACTGCTCTTGGACTTACTTACCCAACTGCTATGCTCTGATGGGGTGGAGTTC 8172  
QY 5942 AAGATCAGGCAAGAGAAAACATGTAAGAAAGTAAATCTCCCAACTCACCAGGATACATG 6001  
Db 8173 AAGATCAGGCAAGAGAAAACATGTAAGAAAGTAAATCTCCCAACTCACCAGGATACATG 8232  
QY 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAATACATGAACCTCCGTAACCC 6061  
Db 8233 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAATACATGAACCTCCGTAACCC 8292  
QY 6062 ATACTGCTGTGTAAGCTTATTTAATACCACTCTCACTGGGCTCCATGAGGTCTCGGCC 6121

Db 8293 ATACTGCGCTGTGTAAGCTTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCC 8352  
QY 6122 AAAACCCCTACTACTGTGTGATATGCTCCCCCTGAACTTTCARGCCATATGTTTCAATCC 6181  
Db 8353 AAAACCCCTACTACTGTGTGATATGCTCCCCCTGAACTTTCARGCCATATGTTTCAATCC 8412  
QY 6182 CTGTACTCTGAACAACTGGAACACTTCCAGCAGAGAAATAACACACTTCCGTTTGTAGTAG 6241  
Db 8413 CTGTACTCTGAACAACTGGAACACTTCCAGCAGAGAAATAACACACTTCCGTTTGTAGTAG 8472  
QY 6242 GACCTCTGTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTTCACTGTGTAAAAATTTA 6301  
Db 8473 GACCTCTGTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTTCACTGTGTAAAAATTTA 8532  
QY 6302 GCAATATCATATACACAACTCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAA 6361  
Db 8533 GCAATATCATATACACAACTCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAA 8592  
QY 6362 TAGTCTGCTTACCTCCTCAGGAATATTTTGTCTGTGTGCTTACCTCAGGCTATCGTTTGTGA 6421  
Db 8593 TAGTCTGCTTACCTCCTCAGGAATATTTTGTCTGTGTGCTTACCTCAGGCTATCGTTTGTGA 8652  
QY 6422 ATGGCTCTTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCCCTCATGRCATCTTACA 6481  
Db 8653 ATGGCTCTTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCCCTCATGRCATCTTACA 8712  
QY 6482 CTGAACAAGATTATATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTAGTACCCATTC 6541  
Db 8713 CTGAACAAGATTATATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTAGTACCCATTC 8772  
QY 6542 TTCTTTTGTATAGAGCAGAGTCTAGTGCTAGTGCTACTAGTACTGGCATTTGGCGGTATCA 6601  
Db 8773 TTCTTTTGTATAGAGCAGAGTCTAGTGCTAGTGCTACTAGTACTGGCATTTGGCGGTATCA 8832  
QY 6602 CCACTCTACTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 6661  
Db 8833 CCACTCTACTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 8892  
QY 6662 TCGCCGACTCCCTGGTGCACCTTGCAGAGTCACTTAACCTCCCTAGCAGAGTAGTCTCTTC 6721  
Db 8893 TCGCCGACTCCCTGGTGCACCTTGCAGAGTCACTTAACCTCCCTAGCAGAGTAGTCTCTTC 8952

## RESULT 16

BD221808  
LOCUS BD221808 10499 bp DNA linear PAT 17-JUL-2003  
DEFINITION Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses.  
ACCESSION BD221808  
VERSION BD221808.1 GI:33031578  
KEYWORDS JP 2002518051-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 10499)  
AUTHORS Alliel, P. M., Perin, J. P. and Rieger, F.  
TITLE Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses  
JOURNAL Patent: JP 2002518051-A 3 25-JUN-2002;  
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE  
COMMENT OS Homo sapiens (human)  
PN JP 2002518051-A/3  
PD 25-JUN-2002  
PF 23-JUN-1999 JP 2000556036  
PR 23-JUN-1998 FR 98/07920  
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC  
C12N15/09, A01K67/027, A61K31/713, A61K39/21, A61K48/00, A61P21/00, PC  
A61P25/00,  
PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC  
sequence and deduced protein sequence family with CC  
human endogenous

CC		retroviral motifs, and their uses	
Key	Location/Qualifiers		
FT	1..10499		
FT	/organism='Homo sapiens (human)'		
FEATURES	Location/Qualifiers		
source	1..10499		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	10.3%;	Score 779;	DB 2; Length 10499;
Best Local Similarity	98.9%;	Pred. No. 0;	
Matches 1899;	Conservative 0;	Mismatches 20;	Indels 1; Gaps 1;
QY	4802	AGGCTTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACATAGTTTCATGAAA	4861
Db	7101	AGGCTTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACATAGTTTCATGAAA	7160
QY	4862	TAAATCCAGATTCGGACTTTCCCGAGGCTTTACAGAGTGACAATAGCCCTTGCTTTCCAGG	4921
Db	7161	TAAATCCAGATTCGGACTTTCCCGAGGCTTTACAGAGTGACAATAGCCCTTGCTTTCCAGG	7220
QY	4922	CCACAGTAAACCCAGGGAGTATCCCGAGGCTTAGGTATACGATATCATCTTACCTGGCGCT	4981
Db	7221	CCACAGTAAACCCAGGGAGTATCCCGAGGCTTAGGTATACGATATCATCTTACCTGGCGCT	7280
QY	4982	GAAGGCCACAGTCTCCTCAGGGAAGGTCCAGAAAATGAATGAAVACTCAAAGGCATCTAA	5041
Db	7281	GAAGGCCACAGTCTCCTCAGGGAAGGTCCAGAAAATGAATGAAVACTCAAAGGCATCTAA	7340
QY	5042	AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGTCGTGTGGCTATAGCCCTTAAAAA	5101
Db	7341	AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGTCGTGTGGCTATAGCCCTTAAAAA	7400
QY	5102	GAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	5161
Db	7401	GAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	7460
QY	5162	CCTTCAATAACCAATGACTTGTGCTTGACCCAGACAGCAACTTAGTTGCGACACATCAC	5221
Db	7461	CCTTCAATAACCAATGACTTGTGCTTGACCCAGACAGCAACTTAGTTGCGACACATCAC	7520
QY	5222	CTCCTTAGCCAAATATCAACAAGTTCCTTAAAAATTAACAAGGAACCTATCCCTGAGAAGA	5281
Db	7521	CTCCTTAGCCAAATATCAACAAGTTCCTTAAAAATTAACAAGGAACCTATCCCTGAGAAGA	7580
QY	5282	GGGAAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	5341
Db	7581	GGGAAAAGAACTATTCCACCCCTTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC	7640
QY	5342	CCCATCCCTAGATACATCTGGGAAGGACCCCTACCCAGTCAATTTTATYTAACCCCACTGC	5401
Db	7641	CCCATCCCTAGATACATCTGGGAAGGACCCCTACCCAGTCAATTTTATYTAACCCCACTGC	7700
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACTTTGAGTCAAACTCTGGATACT	5461
Db	7701	GGTTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACTTTGAGTCAAACTCTGGATACT	7760
QY	5462	GCCAAAGGAACCTGAAAATCCAGGAGCAACGCTAGCTATTTCCTGTGAACCTCTAGAGGA	5521
Db	7761	GCCAAAGGAACCTGAAAATCCAGGAGCAACGCTAGCTATTTCCTGTGAACCTCTAGAGGA	7820
QY	5522	TTTGGCGCTGCTTTCAAAACAAACACAGGAGGAAAGTAACATAAATCAATAATCCCCCA	5581
Db	7821	TTTGGCGCTGCTTTCAAAACAAACACAGGAGGAAAGTAACATAAATCAATAATCCCCCA	7879
QY	5582	TGSGCCTCCCTTATCATATTTTCTKTASTGTTSTTTTACCTCTTTTCACTCTCACTG	5641
Db	7880	TGSGCCTCCCTTATCATATTTTCTKTASTGTTSTTTTACCTCTTTTCACTCTCACTG	7939
QY	5642	CACCCCTCCATGCCGCTGTATGACCAAGTACGCTCCCTTACCMAGAGTTTCTATGGAGAA	5701

DB	7940	CACCCCTCCATGCCCTGTATGACACAGTAGTCCCTTACCAAGAGTTTCTATGAGAA	7999		
QY	5702	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTSTTAAGGAAACCCCA	5761		
DB	8000	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTSTTAAGGAAACCCCA	8059		
QY	5762	CCTTCACTGCCACACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC	5821		
DB	8060	CCTTCACTGCCACACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC	8119		
QY	5822	ATGCAATATCACTATTATTGGACAGAAAATGATTAATCTAGTTGTCTCGAGAGACTTG	5881		
DB	8120	ATGCAATATCACTATTATTGGACAGAAAATGATTAATCTAGTTGTCTCGAGAGACTTG	8179		
QY	5882	GAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTCTGATGGGGTGGAGTTC	5941		
DB	8180	GAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTCTGATGGGGTGGAGTTC	8239		
QY	5942	AAGATCAGGCAAGAGAAAACATGTAAAAGAAATTAATCTCCCAACTCACCCGGGTACATG	6001		
DB	8240	AAGATCAGGCAAGAGAAAACATGTAAAAGAAATTAATCTCCCAACTCACCCGGGTACATG	8299		
QY	6002	GCACCTTAGCCCTTACAAAGGACTTAGATCTCTCAAAAATACATGAAAACCTTCGGTACCC	6061		
DB	8300	GCACCTTAGCCCTTACAAAGGACTTAGATCTCTCAAAAATACATGAAAACCTTCGGTACCC	8359		
QY	6062	ATACTCGCTCGTGAAGCTATTAAATACCACTCAGTGGGCTCATGAGTCTCGGCC	6121		
DB	8360	ATACTCGCTCGTGAAGCTATTAAATACCACTCAGTGGGCTCATGAGTCTCGGCC	8419		
QY	6122	AAACCCCTACTAACTGTTGGATATGCCCTCCCTCGAACTTCARGCCATATGTTTCAATCC	6181		
DB	8420	AAACCCCTACTAACTGTTGGATATGCCCTCCCTCGAACTTCARGCCATATGTTTCAATCC	8479		
QY	6182	CTGTACTCTGAAACAATGGAAACAACCTTCAGCAAGAAATAAACACCACTTCGGTTAGTAG	6241		
DB	8480	CTGTACTCTGAAACAATGGAAACAACCTTCAGCAAGAAATAAACACCACTTCGGTTAGTAG	8539		
QY	6242	GACCTCTGTTTCCNAATSTGGAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTA	6301		
DB	8540	GACCTCTGTTTCCNAATSTGGAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTA	8599		
QY	6302	GCAATACTACATACACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA	6361		
DB	8600	GCAATACTACATACACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA	8659		
QY	6362	TAGTCTGCCTACCTCAGGAATAATTTTGTGTGTGGTACCTCAGCTATCGTTGTGA	6421		
DB	8660	TAGTCTGCCTACCTCAGGAATAATTTTGTGTGTGGTACCTCAGCTATCGTTGTGA	8719		
QY	6422	ATGGCTCTTCAGAACTATGCTTCTCTCATTTCTTAGTGGCCCTCCATATGRCATCTACA	6481		
DB	8720	ATGGCTCTTCAGAACTATGCTTCTCTCATTTCTTAGTGGCCCTCCATATGRCATCTACA	8779		
QY	6482	CTGAAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAAGAGTACCCATTC	6541		
DB	8780	CTGAAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAAGAGTACCCATTC	8839		
QY	6542	TTTCTTTTGTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGGCAATGGCGGTATCA	6601		
DB	8840	TTTCTTTTGTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGGCAATGGCGGTATCA	8899		
QY	6602	CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGACGGG	6661		
DB	8900	CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGACGGG	8959		
QY	6662	TGCGCGACTCCCTGCTCAGCTTGGCAAGATCAACTTAACTCCCTAGCAGTAGTCTCTTC	6721		
DB	8960	TGCGCGACTCCCTGCTCAGCTTGGCAAGATCAACTTAACTCCCTAGCAGTAGTCTCTTC	9019		

LOCUS AR699757 10499 bp DNA linear PAT 14-SEP-2005  
 DEFINITION Sequence 3 from patent US 6919438.  
 ACCESSION AR699757  
 VERSION AR699757.1 GI:75205546  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 10499)  
 AUTHORS Alliel, P.M., Perin, J.-P. and Rieger, F.  
 TITLE Nucleic sequence and deduced protein sequence family with human  
 endogenous retroviral motifs, and their uses  
 JOURNAL Patent: US 6919438-A 3 19-JUL-2005;  
 Institut National de la Sante et de la Recherche Medical-Inserm;  
 Paris;  
 FRX;  
 FEATURES  
 source Location/Qualifiers  
 1. .10499  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Query Match 10.3%; Score 779; DB 2; Length 10499;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 QY 4802 AGCCCTCCCTCTAGGACAGAAAGGCCCAAGAGTAAATAAAGGCACTAGTTCATGAAA 4861  
 DB 7101 AGCCCTCCCTCTAGGACAGAAAGGCCCAAGAGTAAATAAAGGCACTAGTTCATGAAA 7160  
 QY 4862 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTCCAGG 4921  
 DB 7161 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTCCAGG 7220  
 QY 4922 CCACAGTAACCCAGGAGTATCCAGCGCTTAGGTATACGATATCACTTACACTGCGCCT 4981  
 DB 7221 CCACAGTAACCCAGGAGTATCCAGCGCTTAGGTATACGATATCACTTACACTGCGCCT 7280  
 QY 4982 GAAGGCCACAGTCTCAGGGAAGGTGAGAAATGAATGAATGAATGAATGAATGAATGAAT 5041  
 DB 7281 GAAGGCCACAGTCTCAGGGAAGGTGAGAAATGAATGAATGAATGAATGAATGAATGAAT 7340  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTATAGCCCTTAAAAA 5101  
 DB 7341 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTATAGCCCTTAAAAA 7400  
 QY 5102 GAATCTGCAACTTCCCAAAAGAGGAGCTTAGCCCATACGAAATGCTGTATGGAAGGC 5161  
 DB 7401 GAATCTGCAACTTCCCAAAAGAGGAGCTTAGCCCATACGAAATGCTGTATGGAAGGC 7460  
 QY 5162 CCTTCATAACCAATGACCTTGTGCTTGACCAAGACAGCCAACTTAGTTGCGACATCAC 5221  
 DB 7461 CCTTCATAACCAATGACCTTGTGCTTGACCAAGACAGCCAACTTAGTTGCGACATCAC 7520  
 QY 5222 CTCCTTAGCCAAATATCAAGTCTTAAACATTTACAGGAACCTATCCCTCGAAGA 5281  
 DB 7521 CTCCTTAGCCAAATATCAAGTCTTAAACATTTACAGGAACCTATCCCTCGAAGA 7580  
 QY 5282 GGGAAAGAACTATTCACCCWGTGACATGTTAGTCAAGTCCCTTCCTCTAATTC 5341  
 DB 7581 GGGAAAGAACTATTCACCCWGTGACATGTTAGTCAAGTCCCTTCCTCTAATTC 7640  
 QY 5342 CCATCCCTAGATACATCTCGGAAGGACCTACCCAGTCATTTATATACCCCACTGC 5401  
 DB 7641 CCATCCCTAGATACATCTCGGAAGGACCTACCCAGTCATTTATATACCCCACTGC 7700  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGTAGTCAAAATCTCGGATCT 5461  
 DB 7701 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGTAGTCAAAATCTCGGATCT 7760  
 QY 5462 GCCAAGGAACTGAAATCCAGGACACACGCTAGCTATTCTGTGAACCTCTAGAGGA 5521  
 DB 7761 GCCAAGGAACTGAAATCCAGGACACACGCTAGCTATTCTGTGAACCTCTAGAGGA 7820

QY 5522 TTTCGCGCTGCTCTTCAAAACAACACAGAGGAAAGTAACTAAATCATATAATCCCCCA 5581  
 DB 7821 TTTCGCGCTGCTCTTCAAAACAACACAGAGGAAAGTAACTAAATCATATAATCCCCCA 7879  
 QY 5582 TGGSCCTCCCTATCATATTTTCTCTKATSTTSTTTTACCTSTTTTCACTCTCACTG 5641  
 DB 7880 TGGSCCTCCCTATCATATTTTCTCTTACTTCTTTTACCTCTTTTCACTCTCACTG 7939  
 QY 5642 CACCCCTCCATGCCCTGTATGACCACTAGCTCCCTTACCMAGAGTCTTCTATGGAGAA 5701  
 DB 7940 CACCCCTCCATGCCCTGTATGACCACTAGCTCCCTTACCAAGAGTCTTCTATGGAGAA 7999  
 QY 5702 TGACGCTCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCCA 5761  
 DB 8000 TGACGCTCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCCA 8059  
 QY 5762 CTTTCACTGCCACACCCATATGCCCCGCACTGCTATCACTGCACTCTTTGATGC 5821  
 DB 8060 CTTTCACTGCCACACCCATATGCCCCGCACTGCTATCACTGCACTCTTTGATGC 8119  
 QY 5822 ATGCAATATCTCAATTTTGGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTTG 5881  
 DB 8120 ATGCAATATCTCAATTTTGGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTTG 8179  
 QY 5882 GAGTCACTGTCTGTGGACTTACTTCAACCAACTGATGTCTGATGGGGTGGAGTTC 5941  
 DB 8180 GAGTCACTGTCTGTGGACTTACTTCAACCAACTGATGTCTGATGGGGTGGAGTTC 8239  
 QY 5942 AAGATCAGGCAAGAGAAACATGTAAAGAGTATCTCCCACTCACCGGGGTACATG 6001  
 DB 8240 AAGATCAGGCAAGAGAAACATGTAAAGAGTATCTCCCACTCACCGGGGTACATG 8299  
 QY 6002 GCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCC 6061  
 DB 8300 GCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCC 8359  
 QY 6062 ATATCGCTGTGTAAAGCTTATTAATACCACTCTCACTGGGTCTCACTGAGGTCTCGGCC 6121  
 DB 8360 ATATCGCTGTGTAAAGCTTATTAATACCACTCTCACTGGGTCTCACTGAGGTCTCGGCC 8419  
 QY 6122 AAAACCTCTACTAATCTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 6181  
 DB 8420 AAAACCTCTACTAATCTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 8479  
 QY 6182 CTGTACTCAACAAATGGAACAACTTCAGCACAGAAATAAACACCTCCGTTTATAGTAG 6241  
 DB 8480 CTGTACTCAACAAATGGAACAACTTCAGCACAGAAATAAACACCTCCGTTTATAGTAG 8539  
 QY 6242 GACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACTCACTCTGTAAAAATTAA 6301  
 DB 8540 GACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACTCACTCTGTAAAAATTAA 8599  
 QY 6302 GCATATCTACATACACAACTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA 6361  
 DB 8600 GCATATCTACATACACAACTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA 8659  
 QY 6362 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCTATCGTTGTTGA 6421  
 DB 8660 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCTATCGTTGTTGA 8719  
 QY 6422 ATGGCTCTTTCAGAACTATGCTCTCTCTCAATCTTAGTGCCCTCCATGRCATCTACA 6481  
 DB 8720 ATGGCTCTTTCAGAACTATGCTCTCTCTCAATCTTAGTGCCCTCCATGRCATCTACA 8779  
 QY 6482 CTGAACAGATTTATACAGTTATGATATCTAATAGCCCGCAACAAAAGAGTACCCATTC 6541  
 DB 8780 CTGAACAGATTTATACAGTTATGATATCTAATAGCCCGCAACAAAAGAGTACCCATTC 8839  
 QY 6542 TTCTCTTTTGTATAGGACAGGAGTGTAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGT 6601  
 DB 8840 TTCTCTTTTGTATAGGACAGGAGTGTAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGT 8899

QY 6602 CAACCTCTACTCAGTCTTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
Db  
8900 GAACCTCTACTCAGTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 8959  
QY 6662 TCGCCGACTCCCTGGTGCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTGCCTTC 6721  
Db 8960 TCGCCGACTCCCTGGTGCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTGCCTTC 9019

RESULT 18  
AX007980 10499 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 3 from Patent WO9967395.  
ACCESSION AX007980  
VERSION AX007980.1 GI:9995677  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1  
REFERENCE Perin,J.P., Rieger,F. and Alliel,P.M.  
AUTHORS Nucleic sequence and deduced protein sequence family with human  
TITLE endogenous retroviral motifs, and their uses  
JOURNAL Patent: WO 9967395-A 3 29-DEC-1999;  
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER  
FRANCOIS (FR); ALLIEL PATRICK M (FR)  
FEATURES Location/Qualifiers  
source 1..10499  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 10.3%; Score 779; DB 2; Length 10499;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
QY 4802 AGCCCTTCCCTGTAGGACAGAAAGCCCAAGAGTAATAAGGCACCTAGTTCATGAAA 4861  
Db 7101 AGGCCCTTCCCTGTAGGACAGAAAGCCCAAGAGTAATAAGGCACCTAGTTCATGAAA 7160  
QY 4862 TAATCCAGATTCGAGACTTCCCGAGGCTTACAGAGTACAAATAGCCCTGCTTTCAGG 4921  
Db 7161 TAAATCCAGATTCGAGACTTCCCGAGGCTTACAGAGTACAAATAGCCCTGCTTTCAGG 7220  
QY 4922 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 4981  
Db 7221 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 7280  
QY 4982 GAAGGCCACAGTCTCTCAGGGAAGGTGCAGAAAATGAATGAAAYACTCAAAGGACACTTAA 5041  
Db 7281 GAAGGCCACAGTCTCTCAGGGAAGGTGCAGAAAATGAATGAAACACTCAAAGGACACTTAA 7340  
QY 5042 AAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTCGTGCTCCTATAGCCTTAAAAA 5101  
Db 7341 AAAAGCAAACCCAGGAAACCCACCTCACATGGCCTGTCGTGCTCCTATAGCCTTAAAAA 7400  
QY 5102 GAATCTGCAACTTTCCTCCAAAAGAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161  
Db 7401 GAATCTGCAACTTTCCTCCAAAAGAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 7460  
QY 5162 CTTTCATAACCAATGACCTTGTGCTTGACCCAAAGCAGCCAACTTAGTTGGCAGACATCAC 5221  
Db 7461 CTTTCATAACCAATGACCTTGTGCTTGACCCAAAGCAGCCAACTTAGTTGGCAGACATCAC 7520  
QY 5222 CTTTCATAACCAATGACCTTGTGCTTGACCCAAAGCAGCCAACTTAGTTGGCAGACATCAC 5281  
Db 7521 CTTTCATAACCAATGACCTTGTGCTTGACCCAAAGCAGCCAACTTAGTTGGCAGACATCAC 7580  
QY 5282 GGGAAAAGAACTATTCACCCMMGTGCACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 5341

Db 7581 GGGAAAAGAACTATTCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTC 7640  
QY 5342 CCCATCCCTAGATACATCCTGGGAAGGACCTTACCAGTCACTTTTATYTACCCCACTGC 5401  
Db 7641 CCCATCCCTAGATACATCCTGGGAAGGACCTTACCAGTCACTTTTATYTACCCCACTGC 7700  
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGTCAAAATCCTGGATCT 5461  
Db 7701 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGTCAAAATCCTGGATCT 7760  
QY 5462 GCCAAGGAACCTGAAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGA 5521  
Db 7761 GCCAAGGAACCTGAAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGA 7820  
QY 5522 TTTGGCCCTGCTCTTCAAAACAACAACACGAGGAGGAAGTAATAAATCATAAATCCCCA 5581  
Db 7821 TTTGGCCCTGCTCTTCAAAACAACAACGAGGAGGAAGTAATAAATCATAAATCCCCA 7879  
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTAYACCTTSTTTCACTCTCACTG 5641  
Db 7880 TGGCCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCCTTTCCTCTCACTG 7939  
QY 5642 CACCCCTCCATGCGCTGTATGACCACTAGCTCCCTYACOMAGAGTTTCTATGAGAAA 5701  
Db 7940 CACCCCTCCATGCGCTGTATGACCACTAGCTCCCTTACCAAGAGTTTCTATGAGAAA 7999  
QY 5702 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTSTTAAGGAAACCCCA 5761  
Db 8000 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTCTAAGGGAACCCCA 8059  
QY 5762 CTTTCACATGCCACACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 5821  
Db 8060 CTTTCACATGCCACACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 8119  
QY 5822 ATGCAAAATACCTATTATTGGACAGGAAAAATGATTAATCTTAGTTGTCTCTGGAGACTTG 5881  
Db 8120 ATGCAAAATACCTATTATTGGACAGGAAAAATGATTAATCTTAGTTGTCTCTGGAGACTTG 8179  
QY 5882 GAGTCACATGCTGTGGACTTACTTACCCAAACTGGTATGTCATGCGGGGTGGAGTTC 5941  
Db 8180 GAGTCACATGCTGTGGACTTACTTACCCAAACTGGTATGTCATGCGGGGTGGAGTTC 8239  
QY 5942 AAGATCAGCAGAGAAAAACATGTAAGAAAGTAATCTCCCAACTCACCSGGGTACATG 6001  
Db 8240 AAGATCAGCAGAGAAAAACATGTAAGAAAGTAATCTCCCAACTCACCSGGGTACATG 8299  
QY 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCCGTACCC 6061  
Db 8300 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCCGTACCC 8359  
QY 6062 ATACTCGCTGGTAAGCCTATTATAACACCCCTCACTGGGCTCATGAGGTCTCGGCC 6121  
Db 8360 ATACTCGCTGGTAAGCCTATTATAACACCCCTCACTGGGCTCATGAGGTCTCGGCC 8419  
QY 6122 AAAACCCCTACTAACTGTGGATATGCTCCCTCGAATTCARGCCATATGTTTCAATCC 6181  
Db 8420 AAAACCCCTACTAACTGTGGATATGCTCCCTCGAATTCARGCCATATGTTTCAATCC 8479  
QY 6182 CTGTACTCGAACAATGGAACTTTTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAG 6241  
Db 8480 CTGTACTCGAACAATGGAACTTTTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAG 8539  
QY 6242 GACCTCTTGTTCGAATSTGGAAATAACCCATACCTCAAACTGCTGTGTGTAATAATTTA 6301  
Db 8540 GACCTCTTGTTCGAATSTGGAAATAACCCATACCTCAAACTGCTGTGTGTAATAATTTA 8599  
QY 6302 GCAATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCTCTCCACACAAA 6361  
Db 8600 GCAATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCTCTCCACACAAA 8659  
QY 6362 TAGTCTGCTACCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCGTGTGTTGA 6421  
Db 8660 TAGTCTGCTACCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCGTGTGTTGA 8719

```
QY 6422 ATGGCTCTTCAGATCTATGTGCTTCTCTCTCATTTCTAGTCCGCCCCCVATGRCCATCTACA 6481
Db 8720 ATGGCTCTTCAGATCTATGTGCTTCTCTCTCATTTCTAGTCCGCCCTATGACCATCTACA 8779
QY 6482 CTGAACAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTC 6541
Db 8780 CTGAACAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTC 8839
QY 6542 TTCCTTTTGTATAGGAGCAGGAGTGTCTAGTGTGCACTAGGTACTGGCAATGGCGGTATCA 6601
Db 8840 TTCCTTTTGTATAGGAGCAGGAGTGTCTAGTGTGCACTAGGTACTGGCAATGGCGGTATCA 8899
QY 6602 CAACCTCTACTCAGTCTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6661
Db 8900 CAACCTCTACTCAGTCTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 8959
QY 6662 TCGCCGACTCCCTGGTCAACCTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCTTC 6721
Db 8960 TCGCCGACTCCCTGGTCAACCTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCTTC 9019

RESULT 19
AX329572 56093 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 81 from Patent WO0194629.
ACCESSION AX329572
VERSION AX329572.1 GI:18102550
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1
REFERENCE
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 81 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source Location/Qualifiers
1..56093
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 10.3%; Score 779; DB 2; Length 56093;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTCTGAGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 4861
Db 35101 AGGCTTCCCTCTGAGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 35160
QY 4862 TAATTCACAGATTCGGATCTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 4921
Db 35161 TAATTCACAGATTCGGATCTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 35220
QY 4922 CCACAGTAACCCAGGAGTATCCAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 4981
Db 35221 CCACAGTAACCCAGGAGTATCCAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 35280
QY 4982 GAAGGCCACAGTCTCAGGGAAGGTTCGAGAAATGAATGAACTCAAGGACATCTAA 5041
Db 35281 GAAGGCCACAGTCTCAGGGAAGGTTCGAGAAATGAATGAACTCAAGGACATCTAA 35340
QY 5042 AAAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGTCCTATAGCCCTTAAAAA 5101
Db 35341 AAAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGTCCTATAGCCCTTAAAAA 35400
QY 5102 GAATCTGCACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
```

```
Db 35401 GAATCTGCACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 35460
QY 5162 CTTTCTATAACCAATGACCTTGTGACCAAGACAGCAACTTGTAGTTCGAGACATCAC 5221
Db 35461 CTTTCTATAACCAATGACCTTGTGACCAAGACAGCAACTTGTAGTTCGAGACATCAC 35520
QY 5222 CTCCTTAGCCAAATATCAACAGTTCTTAAAAATTAACAAGAACTTATCCCTGAGAAGA 5281
Db 35521 CTCCTTAGCCAAATATCAACAGTTCTTAAAAATTAACAAGAACTTATCCCTGAGAAGA 35580
QY 5282 GGGAAAAGAACTATTCACCCWNGTGACATGGTATTAGTCAAGTCCCTTCCTTAATTC 5341
Db 35581 GGGAAAAGAACTATTCACCCWNGTGACATGGTATTAGTCAAGTCCCTTCCTTAATTC 35640
QY 5342 CCCATCCCTAGATACATCTCGGGAAGGACCTTACCCAGTCAATTTATATACCCCACTGC 5401
Db 35641 CCCATCCCTAGATACATCTCGGGAAGGACCTTACCCAGTCAATTTATATACCCCACTGC 35700
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCTCTGATACT 5461
Db 35701 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCTCTGATACT 35760
QY 5462 GCCAAAGGAACTGAAATCCAGGAGACACGCTAGTATTCTCTGTGAACTCTTAGAGGA 5521
Db 35761 GCCAAAGGAACTGAAATCCAGGAGACACGCTAGTATTCTCTGTGAACTCTTAGAGGA 35820
QY 5522 TTTGCGCTGCTCTTCAAAACAACAACAGGAGAAAGTAACTAAATCATAAATCCCCCA 5581
Db 35821 TTTGCGCTGCTCTTCAAAACAACAACAGGAGAAAGTAACTAAATCATAAATCCCCCA 35879
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTSTTTTCACTCTCACTG 5641
Db 35880 TGGSCCTCCCTTATCATATTTTCTCTKTACTGTCTTTTACCTCTTTTCACTCTCACTG 35939
QY 5642 CACCCCTCATGCCCTGTATGACACAGTACTCCCTTACCMAGAGTTTCTATGAGAA 5701
Db 35940 CACCCCTCATGCCCTGTATGACACAGTACTCCCTTACCMAGAGTTTCTATGAGAA 35999
QY 5702 TGCAGGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761
Db 36000 TGCAGGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 36059
QY 5762 CTTTCACTGCCACACCCATATGCCCGCAACTGTCTATCACTGCACTCTTTTGCATGC 5821
Db 36060 CTTTCACTGCCACACCCATATGCCCGCAACTGTCTATCACTGCACTCTTTTGCATGC 36119
QY 5822 ATGCAAAATCTCATTTATGGAACAGGAAATATGATTAATCTCTAGTCTCTGAGGACTTG 5881
Db 36120 ATGCAAAATCTCATTTATGGAACAGGAAATATGATTAATCTCTAGTCTCTGAGGACTTG 36179
QY 5882 GAGTCACTGCTGTGGACTTACTTCAACCAACTGATGCTGATGGGGTGGAGTTC 5941
Db 36180 GAGTCACTGCTGTGGACTTACTTCAACCAACTGATGCTGATGGGGTGGAGTTC 36239
QY 5942 AAGATCAGCAAGAGAGAAACATGTAAGAGAGTAACTCTCCAACTCACCGGGTACATG 6001
Db 36240 AAGATCAGCAAGAGAGAAACATGTAAGAGAGTAACTCTCCAACTCACCGGGTACATG 36299
QY 6002 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTATCATGAAACCCCTCCGTACCC 6061
Db 36300 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTATCATGAAACCCCTCCGTACCC 36359
QY 6062 ATACTCGCTGCTGTAAGCTTATTAATACCACTGCTGGCTCCATGAGGCTCTCGGCC 6121
Db 36360 ATACTCGCTGCTGTAAGCTTATTAATACCACTGCTGGCTCCATGAGGCTCTCGGCC 36419
QY 6122 AAAACCCCTACTAACTGTGTGATATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATCC 6181
Db 36420 AAAACCCCTACTAACTGTGTGATATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATCC 36479
QY 6182 CTGTACTCTGAAACAATGGAAACAATTTAGCAGCAGAAATAAACCACTCTCCGTTTATAGTAG 6241
```

Db 36480 CTGTACCTGAACAACAACTTCCAGCACAGAAATAAACACCACTTCCTCGTTTTAGTAG 36539

Qy 6242 GACCTCTTTTCCAAATSTGAAATAACCCATACCTCAAACTCCTGTTGTTAAATTTA 6301

Db 36540 GACCTCTTTTCCAAATSTGAAATAACCCATACCTCAAACTCCTGTTGTTAAATTTA 36599

Qy 6302 GCAATACATACACAAACCACTCCCAATGATCAGTGGGTAACTCTCCACACAAA 6361

Db 36600 GCAATACATACACAAACCACTCCCAATGATCAGTGGGTAACTCTCCACACAAA 36659

Qy 6362 TAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA 6421

Db 36660 TAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA 36719

Qy 6422 ATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTCTAGTGCCTCCCTCATGTCATCA 6481

Db 36720 ATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTCTAGTGCCTCCCTCATGTCATCA 36779

Qy 6482 CTGACAAAGATTATACAGTATGTCATATCTAAGCCCGCACAAAGAGTACCCATTC 6541

Db 36780 CTGACAAAGATTATACAGTATGTCATATCTAAGCCCGCACAAAGAGTACCCATTC 36839

Qy 6542 TTCCTTTTGTATAGGAGCAGAGTCTAGGTGCTACTAGGTACTTGGCATTGGCGGTATCA 6601

Db 36840 TTCCTTTTGTATAGGAGCAGAGTCTAGGTGCTACTAGGTACTTGGCATTGGCGGTATCA 36899

Qy 6602 CAACCTCTACTAGTTCTACTACAAATCTCTCAAGAACTAAATGGGGACATGGAACGGG 6661

Db 36900 CAACCTCTACTAGTTCTACTACAAATCTCTCAAGAACTAAATGGGGACATGGAACGGG 36959

Qy 6662 TCGCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGCAGTAGTCCCTTC 6721

Db 36960 TCGCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGCAGTAGTCCCTTC 37019

RESULT 20

HSAC000064

LOCUS Human BAC clone RG083M05 from 7q21-7q22, complete sequence.

DEFINITION Human BAC clone RG083M05 from 7q21-7q22, complete sequence.

ACCESSION AC000064

VERSION AC000064.1 GI:1669369

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 56093)

AUTHORS Pauley, A.

TITLE The sequence of H. sapiens BAC clone RG083M05

JOURNAL Unpublished (1996)

REFERENCE 2 (bases 1 to 56093)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (13-NOV-1996)

COMMENT Department of Genetics, Washington University

St. Louis, MO 63108, USA

e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:

This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).

VECTOR: pBelo

Selection: chloramphenicol

# NEIGHBORING SEQUENCE INFORMATION:

The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H\_RG083M05; actual end is at 56093 of H\_RG083M05

This clone contains STS sKSS1725.

FEATURES	Location/Qualifiers
source	1..56093 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="7" /map="7q21-7q22" /clone="H_RG083M05" /clone_lib="CITB-978SK-B" complement(838..1131) /rpt_family="ALU" c1360..16971
repeat_region	/gene="WUGSC:H_RG083M05.1" join(c1360..1503,4181..4370,4587..4774,6422..6556,9483..9547,11631..11773,11864..12021,13131..13296,14885..14988,16349..16546,16837..16971)
gene	/gene="WUGSC:H_RG083M05.1" /note="ATPase; strong similarity to peroxisome biosynthesis protein PAB1 (P1D.G1172019); coded for by human cDNA C04279 (NID:91467530)" /codon_start=1 /protein_id="AAB46346.1" /db_xref="GI:1669371"
CDS	/translation="KRLNIQKTELVAFSEAVMOPSVLLDLDLIALGPAVPEHESPDFAEQCEILCNVINKKDCDINKFTDLQHVAKETGGFVARDFTVLVDRAHSRLSQISQREKLVLTLDQKALRFLPASLRSVNLKPRDLGWDKIGLHVEVRLMDTIQPAKVCKKKYFELFANLPIRQATGILLYGPPGKTLGAVIARESKMFLSVKPELLSKYIGASEQAVRDIPIRAQAAPCIFFDEFESIAPRRGHDTGTVRVNQLTQLDVGVEIGQVVLAAATSRPDLIDPALLRPGRLDKVCVCPDPQVTVSYLESKQMLHSFLVRLIILNVLSDSLADVDVLOHVASVTSFTGADLKALYNAOLEALHGMLSKMSFLPDESKEFNRYLFGSSVSELSNGTSSDLSSQCLSPASSMOTDLQVPGKDFQSPQPVLRITASQGCQELTQEQDRLADISIKGRYSQSGEDESNNQGPPIKTRIAISQHLMTALGHTRPSTISEDWKNFAEL"
repeat_region	/rpt_family="ALU" complement(4948..5130)
repeat_region	/rpt_family="ALU" complement(6581..7133)
repeat_region	/rpt_family="ALU" complement(7767..8037)
repeat_region	/rpt_family="ALU" complement(8186..8472)
misc_feature	8473..8625 /gene="WUGSC:H_RG083M05.1" /note="match to human 3' EST H75782 (NID:g1049794), bases 287-444"
misc_feature	8841..9161 /gene="WUGSC:H_RG083M05.1" /note="match to human 5' EST H75921 (NID:g1050050), bases 21-348"
misc_feature	9481..9547 /gene="WUGSC:H_RG083M05.1" /note="match to human 5' EST N22627 (NID:g1130501), bases 276-343"
repeat_region	complement(12612..12907) /rpt_family="ALU"
misc_feature	13670..13793 /gene="WUGSC:H_RG083M05.1" /note="match to human 5' EST H41382 (NID:917434), bases 143-266"
repeat_region	13794..13877





Db 35821 TTGGCGCTGCTCTTCAACACCAACACAGGAGAAAGTAATCAAAATCATAAAT-CCCCA 35879  
Qy 5582 TGGSCCTCCCTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCACTG 5641  
Db 35880 TGGCCCTCCCTTATCATATTTTCTCTTTACTGTTCTTTTAACTCCCTCTTTCACTCTCACTG 35939  
Qy 5642 CACCCCTCCATGCGCGCTGTATGACAGTAGTCTCCCTVACCMAGAGTTTCTATGAGAA 5701  
Db 35940 CACCCCTCCATGCGCGCTGTATGACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGAA 35999  
Qy 5702 TGCAGCGTCCCGAAATATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCA 5761  
Db 36000 TGCAGCGTCCCGAAATATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCA 36059  
Qy 5762 CTTTCACTGCCACACCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 5821  
Db 36060 CTTTCACTGCCACACCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 36119  
Qy 5822 ATGCAAACTCATTTATTTGGACAGGAAATGATTAATCTAGTTGTCTGTGGAGACTTG 5881  
Db 36120 ATGCAAACTCATTTATTTGGACAGGAAATGATTAATCTAGTTGTCTGTGGAGACTTG 36179  
Qy 5882 GAGTCACTCTGTGTTGACTTACTTCAACCAACTGGTATGCTGTATGGGGTGGAGTTC 5941  
Db 36180 GAGTCACTCTGTGTTGACTTACTTCAACCAACTGGTATGCTGTATGGGGTGGAGTTC 36239  
Qy 5942 AAGATCAGGCAAGAGAAACATGTAAAGAGTATCTCCCACTCACCCGGGTACATG 6001  
Db 36240 AAGATCAGGCAAGAGAAACATGTAAAGAGTATCTCCCACTCACCCGGGTACATG 36299  
Qy 6002 GCACCTCTAGCCCTACAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC 6061  
Db 36300 GCACCTCTAGCCCTACAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC 36359  
Qy 6062 ATACTCGCTGTGTAAGCTTATTTAATACCAACCTCACTGGGCTCCATGAGTCTGGGCC 6121  
Db 36360 ATACTCGCTGTGTAAGCTTATTTAATACCAACCTCACTGGGCTCCATGAGTCTGGGCC 36419  
Qy 6122 AAAACCTACTAATCTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 6181  
Db 36420 AAAACCTACTAATCTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 36479  
Qy 6182 CTGTACTCAAACTGGAACACTTCAGCACAGAAATTAACCACTTCCTCGTTTGTAGTAG 6241  
Db 36480 CTGTACTCAAACTGGAACACTTCAGCACAGAAATTAACCACTTCCTCGTTTGTAGTAG 36539  
Qy 6242 GACCTCTGTTTCCAATTTGGAATTAACCCATACCTCAAACTCACTGTGTAATAATTTA 6301  
Db 36540 GACCTCTGTTTCCAATTTGGAATTAACCCATACCTCAAACTCACTGTGTAATAATTTA 36599  
Qy 6302 GCAATACTACATACAAACAACTCCCAATGATCAGGTGGGTAATCTCTCCACACAAA 6361  
Db 36600 GCAATACTACATACAAACAACTCCCAATGATCAGGTGGGTAATCTCTCCACACAAA 36659  
Qy 6362 TAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTGTTACCTCAGCTATCGTTGTTCA 6421  
Db 36660 TAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTGTTACCTCAGCTATCGTTGTTCA 36719  
Qy 6422 ATGGCTCTTCAGAACTATGTTGCTCTCTCTCACTTTAGTGGCCCCCATGACATCTACA 6481  
Db 36720 ATGGCTCTTCAGAACTATGTTGCTCTCTCTCACTTTAGTGGCCCCCATGACATCTACA 36779  
Qy 6482 CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCCCGCAACAAAGAGTACCCATTC 6541  
Db 36780 CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCCCGCAACAAAGAGTACCCATTC 36839  
Qy 6542 TTCTCTTTGTTATAGGAGGAGGTGCTAGGTGCATAGTACTGTCATTTGGCGGTATCA 6601  
Db 36840 TTCTCTTTGTTATAGGAGGAGGTGCTAGGTGCATAGTACTGTCATTTGGCGGTATCA 36899  
Qy 6602 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 6661  
Db 36900 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 36959

Qy 6662 TCGCGGACTCCCTGTCACCTTGCAAGATCAACTTAATCTCCCTAGCAGCAGTAGTCCCTTC 6721  
Db 36960 TCGCGGACTCCCTGTCACCTTGCAAGATCAACTTAATCTCCCTAGCAGCAGTAGTCCCTTC 37019

RESULT 21  
AC007566/c 149194 bp DNA linear PRI 01-MAR-2002  
LOCUS Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.  
DEFINITION AC007566  
ACCESSION AC007566.2 GI:11181861  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 149194)  
Sulston, J.E. and Waterston, R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
PUBMED 9847074  
AUTHORS 2 (bases 1 to 149194)  
Du, Z.  
TITLE The sequence of Homo sapiens BAC clone CTB-10G5  
JOURNAL Unpublished (2001)  
REFERENCE 3 (bases 1 to 149194)  
Waterston, R.H.  
AUTHORS Direct Submission  
TITLE Submitted (15-MAY-1999) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 4 (bases 1 to 149194)  
Waterston, R.  
AUTHORS Direct Submission  
TITLE Submitted (02-OCT-2000) Department of Genetics, Washington  
JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE 5 (bases 1 to 149194)  
Waterston, R.H.  
AUTHORS Direct Submission  
TITLE Submitted (16-NOV-2000) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 6 (bases 1 to 149194)  
Waterston, R.H.  
AUTHORS Direct Submission  
TITLE Submitted (03-JAN-2002) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 7 (bases 1 to 149194)  
Waterston, R.H.  
AUTHORS Direct Submission  
TITLE Submitted (06-FEB-2002) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 8 (bases 1 to 149194)  
Waterston, R.  
AUTHORS Direct Submission  
TITLE Submitted (01-MAR-2002) Department of Genetics, Washington  
JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT On Nov 16, 2000 this sequence version replaced gi:4835815.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
----- Summary Statistics  
-----  
Center project name: H\_RG010G05

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping



clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send mail to: [egreen@nhgri.nih.gov](mailto:egreen@nhgri.nih.gov), or see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.regen.com>).  
VECTOR: pBelOAC11  
Selection: chloramphenicol

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

The clone CTB-10G5 contains the entire sequence of CTB-83M5.

#### FEATURES

source	Location/Qualifiers	
	1. .149194	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="7"	
	/map="7q21-7q22"	
	/clone="CTB-10G5"	
	/clone_lib="CTB-978SK-B"	
repeat_region	1. .1634	
	/rpt_family="L1"	
misc_feature	2248. .2388	
	/note="match to EST BG752883 (NID:gl4063536)"	
misc_feature	2248. .2387	
	/note="match to EST AA149693 (NID:gl720635) zn99d06.r1"	
misc_feature	2248. .2387	
	/note="match to EST AW579261 (NID:g7254310)"	
misc_feature	2248. .2387	
	/note="match to EST BG766882 (NID:gl4077535)"	
misc_feature	2248. .2387	
	/note="match to EST BI160365 (NID:gl4620366)"	
misc_feature	2248. .2374	
	/note="similar to Homo sapiens EST BF758865 (NID:gl2106765)"	
misc_feature	2248. .2287	
	/note="match to EST AV686676 (NID:gl0288539)"	
misc_feature	2253. .2387	
	/note="match to EST AU123510 (NID:gl0948226)"	
misc_feature	2344. .2387	
	/note="match to EST AV686676 (NID:gl0288539)"	
repeat_region	2696. .3066	
	/rpt_family="MalR"	
repeat_region	3108. .3392	
	/rpt_family="Alu"	
misc_feature	3540. .3628	
	/note="match to EST AA425526 (NID:gl2106267) zw48b03.r1"	
misc_feature	3540. .3628	
	/note="similar to Mus musculus EST BB253526 (NID:gb946272)"	
misc_feature	3542. .3632	
	/note="match to EST AV686676 (NID:gl0288539)"	
misc_feature	3542. .3628	
	/note="match to EST AA149693 (NID:gl720635) zn99d06.r1"	
misc_feature	3542. .3628	
	/note="match to EST AU123510 (NID:gl0948226)"	
misc_feature	3542. .3628	
	/note="match to EST BG766882 (NID:gl4077535)"	
misc_feature	3542. .3628	
	/note="match to EST BI160365 (NID:gl4620366)"	
misc_feature	3542. .3609	
	/note="match to EST AW579261 (NID:g7254310)"	
misc_feature	3562. .3628	
	/note="match to EST BE272564 (NID:g9146913)"	
misc_feature	3598. .3628	
	/note="match to EST BE299708 (NID:g9183456)"	
misc_feature	3714. .3785	
	/note="match to EST BG260659 (NID:gl2770475)"	
misc_feature	3717. .3785	
	/note="similar to Mus musculus EST BE994936 (NID:gl0678674)"	
misc_feature	3719. .3785	
	/note="match to EST AA149693 (NID:gl720635) zn99d06.r1"	
misc_feature	3719. .3785	
	/note="match to EST AA425526 (NID:gl2106267) zw48b03.r1"	
misc_feature	3719. .3785	
	/note="match to EST AU123510 (NID:gl0948226)"	
misc_feature	3719. .3785	
	/note="match to EST BE272564 (NID:g9146913)"	
misc_feature	3719. .3785	
	/note="match to EST BE299708 (NID:g9183456)"	
misc_feature	3719. .3785	
	/note="match to EST BG766882 (NID:gl4077535)"	
misc_feature	3719. .3785	
	/note="match to EST BI160365 (NID:gl4620366)"	
misc_feature	3719. .3785	
	/note="similar to Mus musculus EST BB253526 (NID:gb946272)"	
misc_feature	3860. .3862	
	/note="match to EST AV686676 (NID:gl0288539)"	
repeat_region	4730. .4827	
	/rpt_family="L2"	
repeat_region	4834. .4925	
	/rpt_family="L1"	
misc_feature	5715. .5860	
	/note="match to EST AA425526 (NID:gl2106267) zw48b03.r1"	
misc_feature	5715. .5860	
	/note="match to EST AU123510 (NID:gl0948226)"	
misc_feature	5715. .5860	
	/note="match to EST BE272564 (NID:g9146913)"	
misc_feature	5715. .5860	
	/note="match to EST BE299708 (NID:g9183456)"	
misc_feature	5715. .5860	
	/note="match to EST BG260659 (NID:gl2770475)"	
misc_feature	5715. .5860	
	/note="match to EST BG766882 (NID:gl4077535)"	
misc_feature	5715. .5860	
	/note="match to EST BI160365 (NID:gl4620366)"	
misc_feature	5715. .5860	
	/note="match to EST BI160365 (NID:gl4620366)"	
Query Match	10.3%; Score 779; DB 5; Length 149194;	
Best Local Similarity	98.9%; Pred. No. 0;	
Matches 1899; Conservative	0; Mismatches 20; Indels 1; Gaps 1;	
QY	4802 AGGCTTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAGGCACCTAGTTCATGAAA	4861
DB	86903 AGGCTTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAGGCACCTAGTTCATGAAA	86844
QY	4862 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCCAGG	4921

Db 86843 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 86784  
Qy 4922 CCACAGTAAACCCAGGAGATATCCCAAGGGTTAGGTATACGATATCACTTACATCGGCCT 4981  
Db 86783 CCACAGTAAACCCAGGAGATATCCCAAGGGTTAGGTATACGATATCACTTACATCGGCCT 86724  
Qy 4982 GAAGGCCACAGTCTCAGGAGGTTCGAGAAATCAATGAAYACTCAAGAGACATCTAA 5041  
Db 86723 GAAGGCCACAGTCTCAGGAGGTTCGAGAAATCAATGAAYACTCAAGAGACATCTAA 86664  
Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGTCTGTGCTATAGCCTTAAAAA 5101  
Db 86663 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGTCTGTGCTATAGCCTTAAAA 86604  
Qy 5102 GAATCTGCAACTTTCCCAAAAGCAGGACTTTCAGCCCATACGAAATGTGTATGGAAGC 5161  
Db 86603 GAATCTGCAACTTTCCCAAAAGCAGGACTTTCAGCCCATACGAAATGTGTATGGAAGC 86544  
Qy 5162 CTTTCATACCAATGACCTTGTGCTTGACCAAGACAGCACTTAGTTGAGACATCAC 5221  
Db 86543 CTTTCATACCAATGACCTTGTGCTTGACCAAGACAGCACTTAGTTGAGACATCAC 86484  
Qy 5222 CTCCTTAGCCAAATCAACAGTTCCTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 5281  
Db 86483 CTCCTTAGCCAAATCAACAGTTCCTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 86424  
Qy 5282 GGGAAAGAACTATTTCCACCCWGTGACATGTGATATAGTCAAGTCCCTTCYCTTAATTC 5341  
Db 86423 GGGAAAGAACTATTTCCACCCWGTGACATGTGATATAGTCAAGTCCCTTCYCTTAATTC 86364  
Qy 5342 CCATCCCTTAGATACATCTCGGAGAGGACCTTACCAGTCACTTTATATACCCCACTGC 5401  
Db 86363 CCATCCCTTAGATACATCTCGGAGAGGACCTTACCAGTCACTTTATATACCCCACTGC 86304  
Qy 5402 GGTAAAGTGGCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGTGATACT 5461  
Db 86303 GGTAAAGTGGCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGTGATACT 86244  
Qy 5462 GCCAAAGAACTCGAAATCCAGGAGACAAACGCTAGTATTCCTGTGAACTCTTAGAGA 5521  
Db 86243 GCCAAAGAACTCGAAATCCAGGAGACAAACGCTAGTATTCCTGTGAACTCTTAGAGA 86184  
Qy 5522 TTTGGCCTGCTCTTCAACAACAACAGGAGGAAAGTAACTAAATCATTAATCCCCA 5581  
Db 86183 TTTGGCCTGCTCTTCAACAACAACAGGAGGAAAGTAACTAAATCATTAATCCCCA 86125  
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTACCCSTTTTCACTCTCACTG 5641  
Db 86124 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTACCCSTTTTCACTCTCACTG 86065  
Qy 5642 CACCCCTCCATCGCGCTGTATGACCAAGTACTCCCTTACCMAGAGTTTCTATGGAGAA 5701  
Db 86064 CACCCCTCCATCGCGCTGTATGACCAAGTACTCCCTTACCAAGAGTTTCTATGGAGAA 86005  
Qy 5702 TGCAGGTCGCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCA 5761  
Db 86004 TGCAGGTCGCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCA 85945  
Qy 5762 CTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCACTCTTTGCGATGC 5821  
Db 85944 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCACTCTTTGCGATGC 85885  
Qy 5822 ATGCAATACTCATTTATGGACAGGAAATGATTAATCCTAGTTGTCTCGAGGACTTG 5881  
Db 85884 ATGCAATACTCATTTATGGACAGGAAATGATTAATCCTAGTTGTCTCGAGGACTTG 85825  
Qy 5882 GAGTCACTCTGTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGTGGAGTTC 5941  
Db 85824 GAGTCACTCTGTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGTGGAGTTC 85765  
Qy 5942 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTATCTCCCAACTCACCCGGGTACATG 6001  
Db 85764 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTATCTCCCAACTCACCCGGGTACATG 85705

Qy 6002 GCACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCC 6061  
Db 85704 GCACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCC 85645  
Qy 6062 ATACTCGCCTCGTAAAGCCTATTAAATACCAACCCCTCACTCGGCTCCATGAGGTCTCGGCC 6121  
Db 85644 ATACTCGCCTCGTAAAGCCTATTAAATACCAACCCCTCACTCGGCTCCATGAGGTCTCGGCC 85585  
Qy 6122 AAAACCCCTACTAACTGTGTGGATATGCTTCCCTCGAATCTCARGCCATATGTTTCAATCC 6181  
Db 85584 AAAACCCCTACTAACTGTGTGGATATGCTTCCCTCGAATCTCARGCCATATGTTTCAATCC 85525  
Qy 6182 CTGTACCTTGAACATGGAACAACTTCAGACAGAAATAAAACCACTTCCGTTTGTAGTAG 6241  
Db 85524 CTGTACCTTGAACATGGAACAACTTCAGACAGAAATAAAACCACTTCCGTTTGTAGTAG 85465  
Qy 6242 GACCTTTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACTGTGTAAAATTTA 6301  
Db 85464 GACCTTTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACTGTGTAAAATTTA 85405  
Qy 6302 GCAATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAA 6361  
Db 85404 GCAATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAA 85345  
Qy 6362 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGA 6421  
Db 85344 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGA 85285  
Qy 6422 ATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTGCCCTCACTATGRCATCTACA 6481  
Db 85284 ATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTGCCCTCACTATGRCATCTACA 85225  
Qy 6482 CTGAACAAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAACAAAAGAGTACCATTTC 6541  
Db 85224 CTGAACAAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAACAAAAGAGTACCATTTC 85165  
Qy 6542 TTCTTTTGTATAGGACAGAGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCA 6601  
Db 85164 TTCTTTTGTATAGGACAGAGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCA 85105  
Qy 6602 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGG 6661  
Db 85104 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGG 85045  
Qy 6662 TCGCGACTCCCTGCTCACTTTCGCAAGATCACTTAACTCCCTAGCAGCAGTAGTCTTTC 6721  
Db 85044 TCGCGACTCCCTGCTCACTTTCGCAAGATCACTTAACTCCCTAGCAGCAGTAGTCTTTC 84985

RESULT 22  
AF156963  
LOCUS  
DEFINITION Homo sapiens human endogenous retrovirus W envC7-1 envelope protein (env) gene, complete cds.  
ACCESSION AF156963  
VERSION AF156963.1 GI:8272467  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1617)  
AUTHORS Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F. and Paranhos-Baccala,G.  
TITLE Chromosomal distribution and coding capacity of the human endogenous retrovirus HERV-W family  
JOURNAL AIDS Res. Hum. Retroviruses 16 (8), 731-740 (2000)  
PUBMED 10826480  
REFERENCE 2 (bases 1 to 1617)  
AUTHORS Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F. and Paranhos-Baccala,G.

TITLE	Direct Submission
JOURNAL	Submitted (09-JUN-1999) UMR103CNRS-bioMerieux, ENS Lyon, 46 allée d'Italie, Lyon 69364, France
FEATURES	Location/Qualifiers
source	1. .1617
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/clones="envC7-1"
gene	/note="human/rodent monochromosomal NIGMS somatic hybrid panel n#2 (Coriell Institute, Camden, NJ, USA); clone GM/NA10791 retaining human chromosome 7"
	1. .1617
	/gene="env"
CDS	1. .1617
	/gene="env"
	/codon_start=1
	/product="envelope protein"
	/protein_id="AAF74215.1"
	/db_xref="GI:8272468"
	/translations="MALPYHIFLTVLLPSFLTAPPCCRMSSSPYQBEFLWMRORP TNCADPSYRSLSGKPTFTATHMPRNCVHSATLCMHANTHYWTGKWINPSCPGGLGV GIVCWTYRTGMSDGGVODQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLRHETLRT HTRLVSLFNTLLTGLHEVSAQNPTNCICLFLNRPYVSPVPEQNNFSTFSTNTTSV LVGLPVLNBEITHSTLSNLTVCVKFNTTYTNSQCIRWPTPTQIVCLPSGIFFCVGTSA YRCLGSSMSKSLFLPVMPTTYTEQDLYSYVISKPRKRVPIPLVFVTCAGVLLTAE TGIGGTTTSTQFYTKLSQELNGDMERVADSLVTLDQLNSLAALVPLNRRLDLLTAE RGCTGLPLGEECCVYNOSGIGIVKEKTEIDRLQRRAEELRNTGPWGLLSQWMPWTL.P FLGPLAAIILLLLFGPCINLNFVSSRIEAVKLQMEPKMQSKTKLYRRPLDRPASP RSDVNDIKGTGPPEISAAQPLLRPNSAGSS"
repeat_region	1. .1617
	/rpt_family="human endogenous retrovirus w"
primer_bind	/rpt_type=dispersed
primer_bind	1. .18
ORIGIN	1597. .1617
	Query Match 9.6%; Score 731; DB 5; Length 1617;
	Best Local Similarity 99.4%; Pred. No. 0;
	Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY	5685 AGATTTCCTATGAGGAATGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db	
	105 AGAGTTTCTATGAGAATGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 164
QY	5745 TTSTAGGGAGACCCACCTTCACTGCCACACCCATATGCCCCCGGACATGCTATCACTC 5804
Db	
	165 TTCTAAGGGAAACCCCACTTCACTGCCACACCCATATGCCCCCGCACTGCTATCACTC 224
QY	5805 TGGCACTCTTTGCATGCATCAAACTACTCATTTTGGACAGGAAAAATGATTAACTCTAG 5864
Db	
	225 TGGCACTCTTTGCATGCATCAAACTACTCATTTTGGACAGGAAAAATGATTAACTCTAG 284
QY	5865 TTGTCTTGGAGGACTTGGAGTCACTGTCTTGGACTTACTTCACCCAAACTGGTATGTC 5924
Db	
	285 TTGTCTCTGGAGGACTTGGAGTCACTGTCTTGGACTTACTTCACCCAAACTGGTATGTC 344
QY	5925 TGATGGGGTGGAGTTCACAGATCAGGCAAGAAAAACATGTAAAAGAGTAATCTCCCA 5984
Db	
	345 TGATGGGGTGGAGTTCAAGATCAGGCAAGAAAAACATGTAAAAGAGTAATCTCCCA 404
QY	5985 ACTCACCGGGTACATGGCACTCTAGCCCTACAAGGACTAGATCTCTCAAAACTACA 6044
Db	
	405 ACTCACCGGGTACATGGCACTCTAGCCCTACAAGGACTAGATCTCTCAAAACTACA 464
QY	6045 TGAACCCCTCCGTACCCATACTCGCCTGGTAAAGCCTATTTTAATACCAACCCCTCACTGGGCT 6104
Db	
	465 TGAACCCCTCCGTACCCATACTCGCCTGGTAAAGCCTATTTTAATACCAACCCCTCACTGGGCT 524
QY	6105 CCATGAGGCTCTCGGCCCAAAACCCCTACTAATCTGTTGGATATGCTCCCTCGAATTCAR 6164
Db	
	525 CCAATGAGGCTCTCGGCCCAAAACCCCTACTAATCTGTTGGATATGCTCCCTCGAATTCAG 584



Db 1003 TTGTCCTGGAGGACTTGGAGTCACTGCTCTGTTGGACTTACTTCAACCAAACTGGTATGTC 1062

Qy 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCA 5984

Db 1063 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCA 1122

Qy 5985 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044

Db 1123 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 1182

Qy 6045 TGAACCTCTCGPACCCATCTCGCTGGTGAAGCTTATTTAATPACCAACCTCACTGGGCT 6104

Db 1183 TGAACCTCTCGPACCCATCTCGCTGGTGAAGCTTATTTAATACCAACCTCACTGGGCT 1242

Qy 6105 CCATGAGTCTCGGCCCAAAACCTACTAATCTGTGGATAGCTCTCCCTGAACCTTCAR 6164

Db 1243 CCATGAGTCTCGGCCCAAAACCTACTAATCTGTGGATAGCTCTCCCTGAACCTTCAG 1302

Qy 6165 GCCATATGTTTCAATCCCTGACTGAAACAATGGAACAACCTTCAGACAGAAATAAACAC 6224

Db 1303 GCCATATGTTTCAATCCCTGACTGAAACAATGGAACAACCTTCAGACAGAAATAAACAC 1362

Qy 6225 CACTCCGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT 6284

Db 1363 CACTCCGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT 1422

Qy 6285 CACTGTGTAAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 6344

Db 1423 CACTGTGTAAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 1482

Qy 6345 AACTCTCCCAACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6404

Db 1483 AACTCTCCCAACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1542

Qy 6405 AGCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAATCTTAGTGCC 6464

Db 1543 AGCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAATCTTAGTGCC 1602

Qy 6465 CCCATGRCATCTACATGAACAAGATTTATACAGTTATGCTATCTAAGCCCGCAA 6524

Db 1603 CCCATGRCATCTACATGAACAAGATTTATACAGTTATGCTATCTAAGCCCGCAA 1662

Qy 6525 CAAAAGTACCAATCTTCTCTTTGTTATAGAGCAGAGTCTAGGTGCATAGGTAC 6584

Db 1663 CAAAAGTACCAATCTTCTCTTTGTTATAGAGCAGAGTCTAGGTGCATAGGTAC 1722

Qy 6585 TGCCATTTGGGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644

Db 1723 TGCCATTTGGGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1782

Qy 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCT 6704

Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCT 1842

Qy 6705 AGCAGCAGTAGTCTCTTC 6721

Db 1843 AGCAGCAGTAGTCTCTTC 1859

RESULT 25

AX000960

LOCUS 1948 bp DNA linear PAT 10-MAR-2000

DEFINITION Sequence 5 from Patent WO902696.

ACCESSION AX000960

VERSION AX000960.1 GI:7241202

KEYWORDS unclassified sequences.

SOURCE unclassified sequences.

ORGANISM unclassified sequences.

REFERENCE 1 (bases 1 to 1948)

AUTHORS Beseme, F. and Blond, J.

TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES OR WITH PREGNANCY DISORDERS

JOURNAL Patent: WO 902696-A 5 21-JAN-1999;  
BIO MERIEUX (FR); BESEME FREDERIC (FR)  
FEATURES Location/Qualifiers  
source 1..1948  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 9.6%; Score 731; DB 2; Length 1948;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5685 AGAGTTTCTATGAGAAATGCGGTCCCGGAAATATGATGCCCAATCGTATAGGAGTCT 5744

Db 823 AGAGTTTCTATGAGAAATGCGGTCCCGGAAATATGATGCCCAATCGTATAGGAGTCT 882

Qy 5745 TTSTAAGGGAACCCCACTTCACCTGCCCAACCCATATGCCCCGCAATGCTATCACTC 5804

Db 883 TTSTAAGGGAACCCCACTTCACCTGCCCAACCCATATGCCCCGCAATGCTATCACTC 942

Qy 5805 TGCCACTCTTTGCATGCATCAATCACTATTATGGAAGGAAATGATTAATCCTAG 5864

Db 943 TGCCACTCTTTGCATGCATCAATCACTATTATGGAAGGAAATGATTAATCCTAG 1002

Qy 5865 TTGCTCTGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGTC 5924

Db 1003 TTGCTCTGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGTC 1062

Qy 5925 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCA 5984

Db 1063 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCA 1122

Qy 5985 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044

Db 1123 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 1182

Qy 6045 TGAACCTCTCGTACCCATCTCGCTGGTGAAGCTTATTTAATPACCAACCTCACTGGGCT 6104

Db 1183 TGAACCTCTCGTACCCATCTCGCTGGTGAAGCTTATTTAATPACCAACCTCACTGGGCT 1242

Qy 6105 CCATGAGTCTCGGCCCAAAACCTCTACTAATCTGTGGATAGCTCTCCCTGAACCTTCAR 6164

Db 1243 CCATGAGTCTCGGCCCAAAACCTCTACTAATCTGTGGATAGCTCTCCCTGAACCTTCAG 1302

Qy 6165 GCCATATGTTTCAATCCCTGACTGAAACAATGGAACAACCTTCAGACAGAAATAAACAC 6224

Db 1303 GCCATATGTTTCAATCCCTGACTGAAACAATGGAACAACCTTCAGACAGAAATAAACAC 1362

Qy 6225 CACTTCCGTTTATAGTAGGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACT 6284

Db 1363 CACTTCCGTTTATAGTAGGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACT 1422

Qy 6285 CACTGTGTAAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 6344

Db 1423 CACTGTGTAAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 1482

Qy 6345 AACTCTCCCAACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6404

Db 1483 AACTCTCCCAACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1542

Qy 6405 AGCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAATCTTAGTGCC 6464

Db 1543 AGCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAATCTTAGTGCC 1602

Qy 6465 CCCATGRCATCTACATGAACAAGATTTATACAGTTATGCTATCTAAGCCCGCAA 6524

Db 1603 CCCATGRCATCTACATGAACAAGATTTATACAGTTATGCTATCTAAGCCCGCAA 1662

Qy 6525 CAAAAGTACCAATCTTCTCTTTGTTATAGAGCAGAGTCTAGGTGCATAGGTAC 6584

Db 1663 CAAAAGTACCAATCTTCTCTTTGTTATAGAGCAGAGTCTAGGTGCATAGGTAC 1722

Qy 6585 TGCCATTTGGGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644

Db 1723 TGCCATTTGGGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1782

Qy 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCT 6704

Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCT 1842

Qy 6705 AGCAGCAGTAGTCTCTTC 6721

Db 1843 AGCAGCAGTAGTCTCTTC 1859

```
QY 6585 TGGCATTGGCGGTATCAAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1723 TGGCATTGGCGGTATCAAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1782
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTTGCAAGATCAACTTAACTCCCT 6704
Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTTGCAAGATCAACTTAACTCCCT 1842
QY 6705 AGCAGCAGTAGTCCTTC 6721
Db 1843 AGCAGCAGTAGTCCTTC 1859

RESULT 26
AX027474
LOCUS AX027474 1948 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 24 from Patent FR2788784.
ACCESSION AX027474
VERSION AX027474.1 GI:10188438
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Mallet,F., Voisset,C. and Paranhos,B.G.
JOURNAL Patent: FR 2788784-A 24 28-JUL-2000;
BIO MERIEUX (FR)
FEATURES
source
1..1948
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 9.6%; Score 731; DB 2; Length 1948;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 823 AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 882
QY 5745 TTSTAAGGAACCCCACTTCACTGCGCCACACCCATATGCGCCGGAAGTCTATCACTC 5804
Db 883 TTSTAAGGAACCCCACTTCACTGCGCCACACCCATATGCGCCGGAAGTCTATCACTC 942
QY 5805 TGCCACTCTTTCATGCATGCAAACTACTATTATGGACAGGAAATGATTAATCCTAG 5864
Db 943 TGCCACTCTTTCATGCATGCAAACTACTATTATGGACAGGAAATGATTAATCCTAG 1002
QY 5865 TTGTCTCGGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGTC 5924
Db 1003 TTGTCTCGGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGTC 1062
QY 5925 TGATGGGGTGGAGTTCAGATCAGCAAGAGAGAAACATGTAAGAGTAACTTCCCA 5984
Db 1063 TGATGGGGTGGAGTTCAGATCAGCAAGAGAGAAACATGTAAGAGTAACTTCCCA 1122
QY 5985 ACTCACCGGTACATGGCAGCTCTAGCCCTACAAAGACTAGATCTCTCAAACTACA 6044
Db 1123 ACTCACCGGTACATGGCAGCTCTAGCCCTACAAAGACTAGATCTCTCAAACTACA 1182
QY 6045 TGAACCCCTCCGTACCCATCTCGCTGTGAAGCCCTATTATTAACCCCTCACTGGGCT 6104
Db 1183 TGAACCCCTCCGTACCCATCTCGCTGTGAAGCCCTATTATTAACCCCTCACTGGGCT 1242
QY 6105 CCATGAGGTCTCGGGCCAAACCCCTACTAACTCTGTTGGATATGCTCCGCCCTGAATTCAR 6164
Db 1243 CCATGAGGTCTCGGGCCAAACCCCTACTAACTCTGTTGGATATGCTCCGCCCTGAATTCAG 1302
QY 6165 GCATATGTTTCAATCCCTGTACCTGAACTGGAACAACTTTCAGCACAGAAATAACAC 6224
```

```
Db 1303 GCCATATGTTTCAATCCCTGTACTGAACTGGAACAACTTCCAGCAGAAATAAACAC 1362
QY 6225 CACTTCCGTTTTAGTAGGACCTCTTGTGTTTCAATSTGGAATAAACCCATACCTCAAACT 6284
Db 1363 CACTTCCGTTTTAGTAGGACCTCTTGTGTTTCAATCTGGAATAAACCCATACCTCAAACT 1422
QY 6285 CACTGTGTAAATTTAGCAATACATACACAACTCCCAATGCAATGCAATGAGTGGGT 6344
Db 1423 CACTGTGTAAATTTAGCAATACATACACAACTCCCAATGCAATGAGTGGGT 1482
QY 6345 AACTCTCTCCACACAAATAGTCTGCTACCTCTCAGGAATATTTTTTGTCTGTGGTACCTC 6404
Db 1483 AACTCTCTCCACACAAATAGTCTGCTACCTCTCAGGAATATTTTTTGTCTGTGGTACCTC 1542
QY 6405 AGCTATCGTTTGTGTAATGGCTCTTCAAGATCTAATGTGCTCTCTCAATCTTGTAGTCC 6464
Db 1543 AGCTATCGTTTGTGTAATGGCTCTTCAAGATCTAATGTGCTCTCTCAATCTTGTAGTCC 1602
QY 6465 CCCATGRCCTACACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 6524
Db 1603 CCCATGRCCTACACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 1662
QY 6525 CAAAAGAGTACCCCATTTCTTCTTTTATAGGAGCAGAGTGTAGTGTGCACTAGGTAC 6584
Db 1663 CAAAAGAGTACCCCATTTCTTCTTTTATAGGAGCAGAGTGTAGTGTGCACTAGGTAC 1722
QY 6585 TGGCAATGGCGGTATCAAACTCTTCACTGTTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1723 TGGCAATGGCGGTATCAAACTCTTCACTGTTTCTACTACAACTATCTCAAGAACTAAA 1782
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGCAGATCAACTTAACTCCCT 6704
Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGCAGATCAACTTAACTCCCT 1842
QY 6705 AGCAGCAGTAGTCCTTC 6721
Db 1843 AGCAGCAGTAGTCCTTC 1859

RESULT 27
AF072505 1948 bp mRNA linear PRI 10-FEB-1999
LOCUS AF072505
DEFINITION Homo sapiens endogenous retrovirus W envelope protein mRNA, partial cds.
ACCESSION AF072505
VERSION AF072505.1 GI:4262289
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H., Mandrand,B. and Mallet,F.
TITLE Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family
JOURNAL J. Virol. 73 (2), 1175-1185 (1999)
PUBMED 9882319
REFERENCE 2 (bases 1 to 1948)
AUTHORS Blond,J.-L., Beseme,F. and Mallet,F.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) UMI03 CNRS-bioMerieux, ENS Lyon, 46 allée d'Italie, Lyon, Cedex 07 69364, France
FEATURES
source
1..1948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl.24.4"
/tissue_type="placenta"
/clone_lib="Clontech 5'-stretch plus library, Cat number
```

```
HL5014a"
repeat_region 1..1948
/tpc_family="HERV-W"
5'UTR <1..6
/nc="R"
7..460
/nc="U5"
/nc="U5"
464..481
/nc="putative primer binding site for tRNA-W"
misc_feature 491..492
/nc="putative splice junction"
719..>1948
/codon_start=1
/product="envelope protein"
/protein_id="AADI454.1"
/db_xref="GI:4262290"
/translators="MALPHYFILFVVSPSFTLTAPPCKRMTSSSPYQBFRLWRMRP
GNIDATYSLSLCKGTFTTAAHPRNCYHSATLCMHANTHYWTGMINPSCPGGLGV
TCWYFTQTDGGGVQDQAREKHKEVISQLTRVHGTSYKGLDLSKLHETLRT
HTRVLSLENTLTLGLHEVSAONPTNCWICLPLNFRPVSTVPEQWNNSTELNITSV
LVGPLVSNLEIHTSLNLCVKFSNTYTNSQIRWVPTQIVLPSGIFVCGTSA
YRCLNASSSMCFSLFSLVPMALYITQDLYSYVISKPRNKRVPILPFGVGLGALG
TGGIGTTSTQFYKLSQELNGDMERVADSLVTLQDLNSLAALVQLNRRALDLTAE
RGGTCFLGEBECYYNQ"
ORIGIN
Query Match 9.6%; Score 731; DB 5; Length 1948;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 5685 AGAGTTTCTATGAGAAATGAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
DB 823 AGAGTTTCTATGAGAAATGAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 882
QY 5745 TTSTAAGGGAAACCCCACTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 5804
DB 883 TTGTAAGGAAACCCCACTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 942
QY 5805 TGCACCTTTTGCATGATGCAATCACTATTAATGGACAGGAAATGATTAACTCTAG 5864
DB 943 TGCACCTTTTGCATGATGCAATCACTATTAATGGACAGGAAATGATTAACTCTAG 1002
QY 5865 TTGCTCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTACCACAACTGGTATGTC 5924
DB 1003 TTGCTCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTACCACAACTGGTATGTC 1062
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTATCCTCA 5984
DB 1063 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTATCCTCA 1122
QY 5985 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 6044
DB 1123 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 1182
QY 6045 TGAACCTCCGTACCATCTGCTGTAAGCCTATTATTAACACCTCTACTGGCT 6104
DB 1183 TGAACCTCCGTACCATCTGCTGTAAGCCTATTATTAACACCTCTACTGGCT 1242
QY 6105 CCATGAGTCTCGGCCCAAAACCTACTACTGTTGATATGCTCCCTGAACTTCAR 6164
DB 1243 CCATGAGTCTCGGCCCAAAACCTACTACTGTTGATATGCTCCCTGAACTTCAR 1302
QY 6165 GCCATATGTTTCAATCCCTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACAC 6224
DB 1303 GCCATATGTTTCAATCCCTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACAC 1362
QY 6225 CACTTCGGTTTATGAGGACTCTTGTTCATSTGGAATAACCCATACCTCAAACT 6284
DB 1363 CACTTCGGTTTATGAGGACTCTTGTTCATSTGGAATAACCCATACCTCAAACT 1422
QY 6285 CACTTCGTAATATTTAGCAATACTACATACACCAACTCCCAATGATCAGGTGGGT 6344
DB 1423 CACTTCGTAATATTTAGCAATACTACATACACCAACTCCCAATGATCAGGTGGGT 1482
```

```
QY 6345 AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGCTACCTC 6404
DB 1483 AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGCTACCTC 1542
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCATCTTAGTGCC 6464
DB 1543 AGCCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCATCTTAGTGCC 1602
QY 6465 CCCATGRCCTACTACACTGAACAAGATTATACAGTTATGTCTATCTAAGCCCGCAA 6524
DB 1603 CCCATGRCCTACTACACTGAACAAGATTATACAGTTATGTCTATCTAAGCCCGCAA 1662
QY 6525 CAAAAGAGTACCATCTCTCTTTTGTATAGAGAGAGAGTCTAGTGCATAGGTAC 6584
DB 1663 CAAAAGAGTACCATCTCTCTTTTGTATAGAGAGAGAGTCTAGTGCATAGGTAC 1722
QY 6585 TGGCATTGGCGGTATCACAACTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAA 6644
DB 1723 TGGCATTGGCGGTATCACAACTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAA 1782
QY 6645 TGGGGACATGGAAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCCT 6704
DB 1783 TGGGGACATGGAAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCCT 1842
QY 6705 AGCAGCAGTAGTCTCTTC 6721
DB 1843 AGCAGCAGTAGTCTCTTC 1859
```

```
RESULT 28
BD221827
LOCUS
DEFINITION
2055 bp DNA linear PAT 17-JUL-2003
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses.
ACCESSION
BD221827
VERSION
BD221827.1 GI:33031597
KEYWORDS
JP 2002518051-A/22.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 2055)
AUTHORS
Allie, P.M., Perin, J.P. and Rieger, F.
TITLE
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL
Patent: JP 2002518051-A 22 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT
OS Homo sapiens (human)
PN JP 2002518051-A/22
PD 25-JUN-2002
PP 23-JUN-1999 JP 2000556036
PR 23-JUN-1998 FR 98/07920
PI PATRICK M ALLIE, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
C12N15/09, A01K67/027, A61K31/711, A61K39/21, A61K48/00, A61P21/00, PC
A61P25/00,
PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT CDS (1)..(2055).
FEATURES
Location/Qualifiers
source
1..2055
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 9.6%; Score 731; DB 2; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```



```
QY 5685 AGAGTTTCTATGGAGATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 495 AGAGTTTCTATGGAGATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 554
QY 5745 TTSTAAGGGAACCCCACTTCACTGCGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804
Db 555 TTCTAAGGGAACCCCACTTCACTGCGCCACACCCCATATGCCCGCAACTGCTATCACTC 614
QY 5805 TGCCACTCTTTGCATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 5864
Db 615 TGCCACTCTTTGCATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 674
QY 5865 TTGTCCTGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 5924
Db 675 TTGTCCTGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 734
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGATTAATCTCCA 5984
Db 735 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGATTAATCTCCA 794
QY 5985 ACTCACCGGGTACATGGACCTCTAGGCCCTTACAAAGACTAGATCTCTCAAAACTACA 6044
Db 795 ACTCACCGGGTACATGGACCTCTAGGCCCTTACAAAGACTAGATCTCTCAAAACTACA 854
QY 6045 TGAAACCCCTCGTACCCATCTCGCTGGTAGCCCTATTATATACACCCCTCACTGGCT 6104
Db 855 TGAAACCCCTCGTACCCATCTCGCTGGTAGCCCTATTATATACACCCCTCACTGGCT 914
QY 6105 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTTCCCTCGAACTTCAR 6164
Db 915 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTTCCCTCGAACTTCAG 974
QY 6165 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCAGAAAAATAAAC 6224
Db 975 GCCATATGTTTCAATCCCTGTACTGAAACAACCTTCAGCAGAAAAATAAAC 1034
QY 6225 CACTTCGGTTTGTAGGACCTTGTTCGAATSTGGNAATAACCCATACCTCAAACT 6284
Db 1035 CACTTCGGTTTGTAGGACCTTGTTCGAATCTGGAAATAACCCATACCTCAAACT 1094
QY 6285 CACTGTGTAAAAATTTAGCAATACATACAAACCAACTCCCAATGATCACTGGGT 6344
Db 1095 CACTGTGTAAAAATTTAGCAATACATACAAACCAACTCCCAATGATCACTGGGT 1154
QY 6345 AACTCCTCCCAACAATAAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404
Db 1155 AACTCCTCCCAACAATAAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1214
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATTTCTAGTGC 6464
Db 1215 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATTTCTAGTGC 1274
QY 6465 CCATATGRCATCTACACTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAA 6524
Db 1275 CCATATGRCATCTACACTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAA 1334
QY 6525 CAAAAAGAGTACCATTCTTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 6584
Db 1335 CAAAAAGAGTACCATTCTTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 1394
QY 6585 TGGCATTTGGCGGTATACAACTCTACTCAGTTCTTACTACAAATCTATCTCAAGAACTAAA 6644
Db 1395 TGGCATTTGGCGGTATACAACTCTACTCAGTTCTTACTACAAATCTATCTCAAGAACTAAA 1454
QY 6645 TGGGACATGGAACGGGTGCGGACTTCCCTGTGTCACTTGGCAAGATCACTTAACTCCCT 6704
Db 1455 TGGGACATGGAACGGGTGCGGACTTCCCTGTGTCACTTGGCAAGATCACTTAACTCCCT 1514
QY 6705 AGCAGCAGTAGTCTTTC 6721
Db 1515 AGCAGCAGTAGTCTTTC 1531
```

```
RESULT 29
AR699776 AR699776 2055 bp DNA linear PAT 14-SEP-2005
LOCUS Sequence 22 from patent US 6919438.
DEFINITION AR699776
ACCESSION AR699776
VERSION AR699776.1 GI:75205567
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2055)
AUTHORS Alliel,P.M., Perin,J.-P. and Rieger,P.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: US 6919438-A 22 19-JUL-2005;
Institut National de la Sante et de la Recherche Medical-Inserm;
Paris;
FRX;
FEATURES
source Location/Qualifiers
1..2055
/mol_type="genomic DNA"
ORIGIN
Query Match 9.6%; Score 731; DB 2; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 5685 AGAGTTTCTATGGAGATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 495 AGAGTTTCTATGGAGATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 554
QY 5745 TTSTAAGGGAACCCCACTTCACTGCGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804
Db 555 TTCTAAGGGAACCCCACTTCACTGCGCCACACCCCATATGCCCGCAACTGCTATCACTC 614
QY 5805 TGCCACTCTTTGCATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 5864
Db 615 TGCCACTCTTTGCATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 674
QY 5865 TTGTCCTGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 5924
Db 675 TTGTCCTGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 734
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGATTAATCTCCA 5984
Db 735 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGATTAATCTCCA 794
QY 5985 ACTCACCGGGTACATGGACCTCTAGGCCCTTACAAAGACTAGATCTCTCAAAACTACA 6044
Db 795 ACTCACCGGGTACATGGACCTCTAGGCCCTTACAAAGACTAGATCTCTCAAAACTACA 854
QY 6045 TGAAACCCCTCGTACCCATCTCGCTGGTAGCCCTATTATATACACCCCTCACTGGCT 6104
Db 855 TGAAACCCCTCGTACCCATCTCGCTGGTAGCCCTATTATATACACCCCTCACTGGCT 914
QY 6105 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTTCCCTCGAACTTCAR 6164
Db 915 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTTCCCTCGAACTTCAG 974
QY 6165 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCAGAAAAATAAAC 6224
Db 975 GCCATATGTTTCAATCCCTGTACTGAAACAACCTTCAGCAGAAAAATAAAC 1034
QY 6225 CACTTCGGTTTGTAGGACCTTGTTCGAATSTGGNAATAACCCATACCTCAAACT 6284
Db 1035 CACTTCGGTTTGTAGGACCTTGTTCGAATCTGGAAATAACCCATACCTCAAACT 1094
QY 6285 CACTGTGTAAAAATTTAGCAATACATACAAACCAACTCCCAATGATCACTGGGT 6344
Db 1095 CACTGTGTAAAAATTTAGCAATACATACAAACCAACTCCCAATGATCACTGGGT 1154
QY 6345 AACTCCTCCCAACAATAAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404
Db 1155 AACTCCTCCCAACAATAAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1214
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATTTCTAGTGC 6464
Db 1215 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATTTCTAGTGC 1274
QY 6465 CCATATGRCATCTACACTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAA 6524
Db 1275 CCATATGRCATCTACACTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAA 1334
QY 6525 CAAAAAGAGTACCATTCTTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 6584
Db 1335 CAAAAAGAGTACCATTCTTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 1394
QY 6585 TGGCATTTGGCGGTATACAACTCTACTCAGTTCTTACTACAAATCTATCTCAAGAACTAAA 6644
Db 1395 TGGCATTTGGCGGTATACAACTCTACTCAGTTCTTACTACAAATCTATCTCAAGAACTAAA 1454
QY 6645 TGGGACATGGAACGGGTGCGGACTTCCCTGTGTCACTTGGCAAGATCACTTAACTCCCT 6704
Db 1455 TGGGACATGGAACGGGTGCGGACTTCCCTGTGTCACTTGGCAAGATCACTTAACTCCCT 1514
QY 6705 AGCAGCAGTAGTCTTTC 6721
Db 1515 AGCAGCAGTAGTCTTTC 1531
```

QY 6345 AACCTCTCCACACAAATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACTCTC 6404  
Db 1155 AACCTCTCCACACAAATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACTCTC 1214  
QY 6405 AGCCTATCGTGTGTGAATGGCTCTTCAGAACTATGTCTCTCTCATTTCTTAGTGCC 6464  
Db 1215 AGCCTATCGTGTGTGAATGGCTCTTCAGAACTATGTCTCTCTCATTTCTTAGTGCC 1274  
QY 6465 CCCATGRCATCTACACTGAAACAGATTTATACAGTTATGTCTATGTCTATAGCCCGCAA 6524  
Db 1275 CCCTATGACCATCTACACTGAAACAGATTTATACAGTTATGTCTATAGCCCGCAA 1334  
QY 6525 CAAAAGAGTACCCATCTTCTCTTTGTATAGGACAGGAGTGTAGTCTATAGCCCGCAA 6584  
Db 1335 CAAAAGAGTACCCATCTTCTCTTTGTATAGGACAGGAGTGTAGTCTATAGCCCGCAA 1394  
QY 6585 TGGCATTTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
Db 1395 TGGCATTTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1454  
QY 6645 TGGGACATGGAACGGGTCCGCACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCT 6704  
Db 1455 TGGGACATGGAACGGGTCCGCACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCT 1514  
QY 6705 AGCAGCAGTAGTCTCTC 6721  
Db 1515 AGCAGCAGTAGTCTCTC 1531

## RESULT 30

AX007999  
LOCUS AX007999 2055 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 22 from Patent WO9967395.  
ACCESSION AX007999  
VERSION AX007999.1 GI:9995696

## KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.

## REFERENCE

1 Perin, J.P., Rieger, F. and Alliel, P.M.  
Nucleic sequence and deduced protein sequence family with human  
endogenous retroviral motifs, and their uses  
Patent: WO 9967395-A 22 DEC-1999;  
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER  
FRANCOIS (FR); ALLIEL PATRICK M (FR)

## FEATURES

source  
1..2055  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
CDS  
1..87  
/note="unassigned protein product"  
/codon\_start=1  
/protein\_id="CAC07392.1"  
/db\_xref="GI:9995697"  
/translation="PKTNLVADITSLAKYQVVLKTLQGYTP"

## ORIGIN

Query Match 9.6%; Score 731; DB 2; Length 2055;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 5685 AGAGTTTCTATGAGAAATGAGCGTCCGGAATATTTATGCCCCTATAGGAGTCT 5744  
Db 495 AGAGTTTCTATGAGAAATGAGCGTCCGGAATATTTATGCCCCTATAGGAGTCT 554  
QY 5745 TTSTAAGGGAACCCACCTTCACTGCCACCCATATGCCCGCAACTGCTATCACTC 5804  
Db 555 TTSTAAGGGAACCCACCTTCACTGCCACCCATATGCCCGCAACTGCTATCACTC 614

Search completed: August 2, 2006, 22:56:40  
Job time : 27465 secs

QY 5805 TGGCACTCTTTGCATGCATGCAAAATACTCATTATTGGACAGGAAAAATGATTAACTCCTAG 5864  
Db 615 TGGCACTCTTTGCATGCATGCAAAATACTCATTATTGGACAGGAAAAATGATTAACTCCTAG 674  
QY 5865 TTGTCTCTGGAGGACTTTGGAGTCACTCTCTGTGGACTTACTTCCACCAACTGGTATGTC 5924  
Db 675 TTGTCTCTGGAGGACTTTGGAGTCACTCTCTGTGGACTTACTTCCACCAACTGGTATGTC 734  
QY 5925 TGATGGGGTGGAGTTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAACTAACTCCCA 5984  
Db 735 TGATGGGGTGGAGTTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAACTAACTCCCA 794  
QY 5985 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACA 6044  
Db 795 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACA 854  
QY 6045 TGAACCCCTCCGTACCATCTACCTCGCTGGTAAAGCTTATTTAATACCACTCTCACTGGCT 6104  
Db 855 TGAACCCCTCCGTACCATCTACCTCGCTGGTAAAGCTTATTTAATACCACTCTCACTGGCT 914  
QY 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAATCTGTGGATATGCTCCCTCGAACTTTCA 6164  
Db 915 CCATGAGGTCTCGGCCCAAAACCCCTACTAATCTGTGGATATGCTCCCTCGAACTTTCA 974  
QY 6165 GCCATATGTTTCATCCCTGTACTGAAACAATGGAAACAACCTTCAGCAGCAAAATAACAC 6224  
Db 975 GCCATATGTTTCATCCCTGTACTGAAACAATGGAAACAACCTTCAGCAGCAAAATAACAC 1034  
QY 6225 CACTTCGGTTTAGTAGGACCTCTGTTCCTCAATSTGGAATAACCCATACCTCAACCT 6284  
Db 1035 CACTTCGGTTTAGTAGGACCTCTGTTCCTCAATSTGGAATAACCCATACCTCAACCT 1094  
QY 6285 CACTCTGTAAAAATTTAGCAATACTACATACAAACCACTCCCAATGATCAGGTGGGT 6344  
Db 1095 CACTCTGTAAAAATTTAGCAATACTACATACAAACCACTCCCAATGATCAGGTGGGT 1154  
QY 6345 AACTCTCCCAACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTC 6404  
Db 1155 AACTCTCCCAACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTC 1214  
QY 6405 AGCCTATCGTGTGTGAATGGCTCTTCAGAACTATGTCTCTCTCTCATTTCTTAGTGCC 6464  
Db 1215 AGCCTATCGTGTGTGAATGGCTCTTCAGAACTATGTCTCTCTCTCATTTCTTAGTGCC 1274  
QY 6465 CCYATGRCATCTACACTGAAACAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 6524  
Db 1275 CCCTATGACCATCTACACTGAAACAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 1334  
QY 6525 CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACAGGAGTGTAGGTGCACTAGGTAC 6584  
Db 1335 CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACAGGAGTGTAGGTGCACTAGGTAC 1394  
QY 6585 TGGCATTGGCGGTATCAGAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
Db 1395 TGGCATTGGCGGTATCAGAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1454  
QY 6645 TGGGACATGGAACGGGTCCGCACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCT 6704  
Db 1455 TGGGACATGGAACGGGTCCGCACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCT 1514  
QY 6705 AGCAGCAGTAGTCTCTC 6721  
Db 1515 AGCAGCAGTAGTCTCTC 1531

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:41 ; Search time 1251 Seconds  
(without alignments)  
11340.325 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caacaatcggtataaac.....tattaaatctgcartgcr 7582

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1403666 seqs, 935554401 residues

Word size : 6

Total number of hits satisfying chosen parameters: 2781113

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : Issued Patents NA:\*

- 1: /EMC\_Celerra\_SID33/ptodata/2/ina/1\_COMB.seq:\*
- 2: /EMC\_Celerra\_SID33/ptodata/2/ina/5\_COMB.seq:\*
- 3: /EMC\_Celerra\_SID33/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /EMC\_Celerra\_SID33/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /EMC\_Celerra\_SID33/ptodata/2/ina/7\_COMB.seq:\*
- 6: /EMC\_Celerra\_SID33/ptodata/2/ina/H\_COMB.seq:\*
- 7: /EMC\_Celerra\_SID33/ptodata/2/ina/PCUTUS\_COMB.seq:\*
- 8: /EMC\_Celerra\_SID33/ptodata/2/ina/PP\_COMB.seq:\*
- 9: /EMC\_Celerra\_SID33/ptodata/2/ina/RE\_COMB.seq:\*
- 10: /EMC\_Celerra\_SID33/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	779	10.3	2599	US-09-719-554-1	Sequence 1, Appli
2	779	10.3	8523	US-09-573-080A-21	Sequence 21, Appl
3	779	10.3	8523	US-09-854-867-21	Sequence 21, Appl
4	779	10.3	10499	US-09-719-554-3	Sequence 3, Appli
5	779	10.3	13537	US-09-949-016-17382	Sequence 22, Appl
6	731	9.6	2055	US-09-719-554-22	Sequence 5640, Ap
7	731	9.6	2763	US-09-949-016-5640	Sequence 3, Appli
8	680	9.0	2946	US-09-175-928-3	Sequence 3, Appli
9	267	3.5	1326	US-09-719-554-2	Sequence 2, Appli
10	211	2.8	711	US-09-719-554-20	Sequence 20, Appl
11	190	2.5	542	US-08-686-878A-48	Sequence 48, Appl
12	190	2.5	542	US-08-721-489-2	Sequence 28, Appl
13	183	2.4	1080	US-09-719-554-28	Sequence 28, Appl
14	149	2.0	443	US-09-513-999C-21394	Sequence 21394, A
15	89	1.2	711	US-09-719-554-21	Sequence 21, Appl
16	76	1.0	46340	US-09-719-554-66	Sequence 66, Appl
17	73	1.0	1481	US-08-979-847B-105	Sequence 105, App
18	64	0.8	168394	US-09-949-016-13002	Sequence 13002, A
19	63	0.8	1799	US-09-719-554-5	Sequence 5, Appli
20	63	0.8	145320	US-09-949-016-15858	Sequence 15858, A
21	58	0.8	758	US-08-979-847B-112	Sequence 112, App
22	56	0.7	77772	US-09-949-016-17417	Sequence 17417, A
23	56	0.7	77997	US-09-949-016-12249	Sequence 12249, A

#### ALIGNMENTS

RESULT 1

55	0.7	2349	3	US-10-104-047-1148	Sequence 1148, Ap
55	0.7	149971	3	US-09-949-016-13590	Sequence 13590, A
54	0.7	1511	3	US-08-979-847B-117	Sequence 117, App
53	0.7	286	3	US-09-403-343B-32	Sequence 32, Appl
53	0.7	693	3	US-08-691-563C-88	Sequence 88, Appl
53	0.7	693	3	US-09-374-766-88	Sequence 88, Appl
53	0.7	693	3	US-08-979-847B-82	Sequence 82, Appl
53	0.7	1196	3	US-08-691-563C-56	Sequence 56, Appl
53	0.7	1196	3	US-09-374-766-56	Sequence 56, Appl
53	0.7	1196	3	US-08-979-847B-52	Sequence 52, Appl
53	0.7	1216	3	US-09-719-554-7	Sequence 7, Appli
53	0.7	1577	3	US-08-691-563C-89	Sequence 89, Appl
53	0.7	1577	3	US-09-374-766-89	Sequence 89, Appl
53	0.7	1577	3	US-08-979-847B-83	Sequence 83, Appl
53	0.7	2784	3	US-09-719-554-4	Sequence 4, Appli
51	0.7	601	3	US-09-949-016-202253	Sequence 202253, A
50	0.7	279	2	US-08-686-878A-50	Sequence 50, Appl
50	0.7	279	2	US-08-721-489-4	Sequence 4, Appli
49	0.6	601	3	US-09-949-016-133739	Sequence 133739, A
49	0.6	635	3	US-08-979-847B-102	Sequence 102, App
49	0.6	1774	3	US-09-719-554-17	Sequence 17, Appl
49	0.6	245286	3	US-09-949-016-15497	Sequence 15497, A
47	0.6	773	3	US-09-719-554-67	Sequence 67, Appl
47	0.6	1489	3	US-09-719-554-6	Sequence 6, Appli
46	0.6	601	3	US-09-949-016-202370	Sequence 202370, A
46	0.6	84571	3	US-09-949-016-17420	Sequence 17420, A
44	0.6	764	3	US-08-979-847B-126	Sequence 126, App
44	0.6	800	3	US-08-979-847B-127	Sequence 127, App
44	0.6	1600	3	US-08-979-847B-206	Sequence 206, App
44	0.6	1600	3	US-08-979-847B-207	Sequence 207, App
43	0.6	506	3	US-09-621-976-2599	Sequence 2599, App
43	0.6	601	3	US-09-949-016-44297	Sequence 44297, A
43	0.6	601	3	US-09-949-016-202252	Sequence 202252, A
43	0.6	1740	3	US-09-719-554-61	Sequence 61, Appl
43	0.6	2364	3	US-08-979-847B-88	Sequence 88, Appl
43	0.6	2391	3	US-08-691-563C-57	Sequence 57, Appl
43	0.6	2391	3	US-09-374-766-57	Sequence 57, Appl
43	0.6	2391	3	US-08-979-847B-53	Sequence 53, Appl
43	0.6	61108	3	US-09-949-002-608	Sequence 608, App
41	0.5	601	3	US-09-949-016-44296	Sequence 44296, A
41	0.5	601	3	US-09-949-016-202250	Sequence 202250, A
41	0.5	601	3	US-09-949-016-202251	Sequence 202251, A
41	0.5	1329	3	US-08-979-847B-108	Sequence 108, App
40	0.5	601	3	US-09-949-016-146560	Sequence 146560, A
40	0.5	601	3	US-09-949-016-146561	Sequence 146561, A
40	0.5	601	3	US-09-949-016-146564	Sequence 146564, A
40	0.5	601	3	US-09-949-016-146565	Sequence 146565, A
40	0.5	780	3	US-09-573-080A-385	Sequence 385, App
40	0.5	780	3	US-09-854-867-385	Sequence 385, App
40	0.5	1597	3	US-08-979-847B-205	Sequence 205, App
40	0.5	2304	3	US-08-979-847B-87	Sequence 87, Appl
39	0.5	601	3	US-09-949-016-44298	Sequence 44298, A
39	0.5	979	3	US-09-719-554-16	Sequence 16, Appl
38	0.5	85	3	US-08-979-847B-195	Sequence 195, App
38	0.5	410	4	US-09-880-107-538	Sequence 538, App
38	0.5	645	2	US-08-471-724-8	Sequence 8, Appli
38	0.5	645	2	US-08-471-969-8	Sequence 8, Appli
38	0.5	645	2	US-08-384-137-8	Sequence 8, Appli
38	0.5	645	2	US-08-470-006A-8	Sequence 8, Appli
38	0.5	645	3	US-08-691-563C-8	Sequence 8, Appli
38	0.5	645	3	US-09-200-990-8	Sequence 8, Appli
38	0.5	645	3	US-09-133-411-8	Sequence 8, Appli
38	0.5	645	3	US-09-374-766-8	Sequence 8, Appli
38	0.5	645	3	US-08-979-847B-8	Sequence 8, Appli
38	0.5	1158	2	US-08-471-724-1	Sequence 1, Appli
38	0.5	1158	2	US-08-471-969-1	Sequence 1, Appli
38	0.5	1158	2	US-08-384-137-1	Sequence 1, Appli

[illegible]

QY 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
 Db 1950 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 2009  
 QY 6662 TCGCCGACTCCCTGGTCACTTTCAGAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 6721  
 Db 2010 TCGCCGACTCCCTGGTCACTTTCAGAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 2069

RESULT 2

US-09-573-080A-21  
 ; Sequence 21, Application US/09573080A  
 ; Patent No. 6828097  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOAN, KNOLL  
 ; APPLICANT: ROGAN, PETER  
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATION  
 ; FILE REFERENCE: 30307  
 ; CURRENT APPLICATION NUMBER: US/09/573,080A  
 ; CURRENT FILING DATE: 2000-05-16  
 ; NUMBER OF SEQ ID NOS: 479  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 21  
 ; LENGTH: 8523  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: repeat region  
 ; LOCATION: (1)..(8523)  
 ; OTHER INFORMATION: herv17  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A  
 ; TITLE: Prototypic sequences for human repetitive DNA  
 ; JOURNAL: Journal of Molecular Evolution  
 ; VOLUME: 35  
 ; ISSUE: 4  
 ; PAGES: 286-291  
 ; DATE: 1992-10-  
 ; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
 ; DATABASE ENTRY DATE: 1996-01-26  
 ; DATABASE ENTRY DATE: 1996-01-26  
 US-09-573-080A-21

Query Match 10.3%; Score 779; DB 3; Length 8523;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 QY 4802 AGSCCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAATAAGGCACTAGTTTCATGAAA 4861  
 Db 6071 AGSCCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAATAAGGCACTAGTTTCATGAAA 6130  
 QY 4862 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 4921  
 Db 6131 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 6190  
 QY 4922 CCACAGTAACCCAGGAGGATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCT 4981  
 Db 6191 CCACAGTAACCCAGGAGGATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCT 6250  
 QY 4982 GAAGGCCACAGTCTCAGGAGGAGGTGAGAAATGAATGAAYACTCAAAAGGACATCTAA 5041  
 Db 6251 GAAGGCCACAGTCTCAGGAGGAGGTGAGAAATGAATGAAYACTCAAAAGGACATCTAA 6310  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGCTGTGCTATAGCTTAAATA 5101  
 Db 6311 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGCTGTGCTATAGCTTAAATA 6370  
 QY 5102 GAATCTGCAACTTCCCAAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161  
 Db 6371 GAATCTGCAACTTCCCAAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 6430  
 QY 5162 CCTTCATAACCAATGACCTTGTGCTTGAACCAAGACAGCCAACTTAGTTAGTGCAGACATCAC 5221

Db 6431 CCTTCATAACCAATGACCTTGTGCTTGAACCAAGACAGCCAACTTAGTTAGCAGACATCAC 6490  
 QY 5222 CTCTTTAGGCAAAATATCAACAAAGTTCTTAAACACATTACAAGGAACCTATCCCTCGAGAGA 5281  
 Db 6491 CTCTTTAGGCAAAATATCAACAAAGTTCTTAAACACATTACAAGGAACCTATCCCTCGAGAGA 6550  
 QY 5282 GGGAAAAGAACTATTCACCCGCGGTCACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 5341  
 Db 6551 GGGAAAAGAACTATTCACCCGCTTGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 6610  
 QY 5342 CCATCCCTAGATACATCTCTGGAGGACCCCTACCCAGTCACTTTATATATACCCCAACTGC 5401  
 Db 6611 CCATCCCTAGATACATCTCTGGAGGACCCCTACCCAGTCACTTTATATACCCCAACTGC 6670  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATACT 5461  
 Db 6671 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATACT 6730  
 QY 5462 GCCAAAGGAACCTGAAAATCCAGGAGCAACCGTAGCTATTCTGTGAACTCTAGAGGA 5521  
 Db 6731 GCCAAAGGAACCTGAAAATCCAGGAGCAACCGTAGCTATTCTGTGAACTCTAGAGGA 6790  
 QY 5522 TTTGGCCTGCTCTTCAACAAACCAACAGGAGGAAAGTAACATAAATCAATCAATCCCCA 5581  
 Db 6791 TTTGGCCTGCTCTTCAACAAACCAACAGGAGGAAAGTAACATAAATCAATCAATCCCCA 6849  
 QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCTTTTACCTCTCCTG 5641  
 Db 6850 TGGSCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCTCTTCTCCTG 6909  
 QY 5642 CACCCCTCCATCGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAA 5701  
 Db 6910 CACCCCTCCATCGCGCTGTATGACAGTAGTCCCTTACCAAGAGTTTCTATGAGAA 6969  
 QY 5702 TGCAGGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCA 5761  
 Db 6970 TGCAGGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCA 7029  
 QY 5762 CTTTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTCTGCACTCTTTGATGC 5821  
 Db 7030 CTTTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTCTGCACTCTTTGATGC 7089  
 QY 5822 ATGCAAAATCTCATTTATGGAACAGAAATAATGATTAATCTAGTTGTCTCTGAGGAGCTTG 5881  
 Db 7090 ATGCAAAATCTCATTTATGGAACAGAAATAATGATTAATCTAGTTGTCTCTGAGGAGCTTG 7149  
 QY 5882 GAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGGTGGAGTTC 5941  
 Db 7150 GAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGGTGGAGTTC 7209  
 QY 5942 AAGATCAGGCAAGAGAAACCATGTAAAGAAAGTAAATCTCCCAACTCACCSGGGTACATG 6001  
 Db 7210 AAGATCAGGCAAGAGAAACCATGTAAAGAAAGTAAATCTCCCAACTCACCSGGGTACATG 7269  
 QY 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 6061  
 Db 7270 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 7329  
 QY 6062 ATACTGCGCTGTGTAAAGCTTATTAATAACCCCTCACTGGGCTCCATGAGGTCTCGGCC 6121  
 Db 7330 ATACTGCGCTGTGTAAAGCTTATTAATAACCCCTCACTGGGCTCCATGAGGTCTCGGCC 7389  
 QY 6122 AAAACCCCTACTTAACCTGTTGGATATGCTCCCTGAACTTTCARGGCATATGTTCAATCC 6181  
 Db 7390 AAAACCCCTACTTAACCTGTTGGATATGCTCCCTGAACTTTCARGGCATATGTTCAATCC 7449  
 QY 6182 CTGTACCTGAAACAAATGGAAACAACTTTCAGCAGAAATAAACCACCTTCCGTTTTAGTAG 6241  
 Db 7450 CTGTACCTGAAACAAATGGAAACAACTTTCAGCAGAAATAAACCACCTTCCGTTTTAGTAG 7509  
 QY 6242 GACCTCTGTTTCCCAATSTGGAAATAACCCATACCTCAAACTCAACCTCTGTGTAAATTTA 6301

```
Db 7510 GACCTCTTTTCCAATCTGGAATAATACCCATACCTCAAACTCACCTGCTGTGTAATAATTTA 7569
Qy 6302 GCAATACCTACATACACAACTCCCAATGCATCAGGTGGTAACTCCTCCACACAAA 6361
Db 7570 GCAATACCTACATACACAACTCCCAATGCATCAGGTGGTAACTCCTCCACACAAA 7629
Qy 6362 TAGTCTGCTACCTCCAGGATATTTTGTCTGTGTGTTACCTCAGCCTATCGTTGTTGA 6421
Db 7630 TAGTCTGCTACCTCCAGGATATTTTGTCTGTGTGTTACCTCAGCCTATCGTTGTTGA 7689
Qy 6422 ATGGCTTTTCAAGATCTATGCTCTCTCTCATTTCTTAGTGCCCCYATGCCATCTACA 6481
Db 7690 ATGGCTTTTCAAGATCTATGCTCTCTCTCATTTCTTAGTGCCCCYATGCCATCTACA 7749
Qy 6482 CTGAACAAGATTTATACAGTTATGTGATATCTAAGCCCGGCAACAAAAGAGTACCATTTC 6541
Db 7750 CTGAACAAGATTTATACAGTTATGTGATATCTAAGCCCGGCAACAAAAGAGTACCATTTC 7809
Qy 6542 TTCCCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATCA 6601
Db 7810 TTCCCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATCA 7869
Qy 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 6661
Db 7870 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 7929
Qy 6662 TGGCGGACTCTCTGCTCAGCTTGGCAAGATCAACTTAATCTCCTAGCAGAGTAGTCCCTTC 6721
Db 7930 TGGCGGACTCTCTGCTCAGCTTGGCAAGATCAACTTAATCTCCTAGCAGAGTAGTCCCTTC 7989

RESULT 3
US-09-854-867-21
; Sequence 21, Application US/09854867
; Patent No. 7014997
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: herv17
US-09-854-867-21

Query Match 10.3%; Score 779; DB 5; Length 8523;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

Qy 4802 AGGCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCAATGAAA 4861
Db 6071 AGGCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCAATGAAA 6130
Qy 4862 TAAATCCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGG 4921
Db 6131 TAAATCCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGG 6190
Qy 4922 CCACAGTAACCCAGGAGTATCCCGAGGCTTAGGTATACGATATCACTTACACTGGCCT 4981
Db 6191 CCACAGTAACCCAGGAGTATCCCGAGGCTTAGGTATACGATATCACTTACACTGGCCT 6250
Qy 4982 GAAAGCCACAGTCTCAGGGAAGGTCCGAGAAAATGAATGAAAYACTCAAGAGGACATCTAA 5041
Db 6251 GAAAGCCACAGTCTCAGGGAAGGTCCGAGAAAATGAATGAAAYACTCAAGAGGACATCTAA 6310
```

```
Qy 5042 AAAAGCAAAACCCAGGAACCCACCTCACAATGGCCTGTCTCTTGGCTATAGCCTTAAAAA 5101
Db 6311 AAAAGCAAAACCCAGGAACCCACCTCACAATGGCCTGTCTCTTGGCTATAGCCTTAAAAA 6370
Qy 5102 GAAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATAGCAAAATGCTGTATGGAAGGC 5161
Db 6371 GAAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATAGCAAAATGCTGTATGGAAGGC 6430
Qy 5162 CTTTCATAACCAATGACCTTGTGTGACCCAAAGACGCCAATCTTAGTTGAGAGACATCAC 5221
Db 6431 CTTTCATAACCAATGACCTTGTGTGACCCAAAGACGCCAATCTTAGTTGAGAGACATCAC 6490
Qy 5222 CTCCTTAGCCAAATATCAAGAGTTCTTAAACATTACAGGAACTTATCCCTGAGAGA 5281
Db 6491 CTCCTTAGCCAAATATCAAGAGTTCTTAAACATTACAGGAACTTATCCCTGAGAGA 6550
Qy 5282 GGGAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCTCTAAATTC 5341
Db 6551 GGGAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCTCTAAATTC 6610
Qy 5342 CCCATCCCTAGATACATCTCTGGGAAGGACCCTACCCAGTCAATTTATYTAACCCCAACTGC 5401
Db 6611 CCCATCCCTAGATACATCTCTGGGAAGGACCCTACCCAGTCAATTTATCTACCCCAACTGC 6670
Qy 5402 GGTAAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGATCAATCTCTGATACT 5461
Db 6671 GGTAAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGATCAATCTCTGATACT 6730
Qy 5462 GCCAAAAGAACTGAAAATCCAGAGACCAACGCTAGCTATTCTCTGTAACCTCTAGAGA 5521
Db 6731 GCCAAAAGAACTGAAAATCCAGAGACCAACGCTAGCTATTCTCTGTAACCTCTAGAGA 6790
Qy 5522 TTTGCGCTGCTCTTCAAAACAAACACGAGGAGAAAGTAACTTAAATCATAAATCCCCCA 5581
Db 6791 TTTGCGCTGCTCTTCAAAACAAACACGAGGAGAAAGTAACTTAAATCATAAATCCCCCA 6849
Qy 5582 TGGSGCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTAACTCTTCTCTCTCTCTG 5641
Db 6850 TGGSGCTCCCTTATCATATTTTCTCTKTACTGTTTCTTTACCTCTTCTCTCTCTG 6909
Qy 5642 CACCCCTCCATGCGCTGTATGACCACTGCTCCCTTACCMAGAGTTTCTTATGAGAAA 5701
Db 6910 CACCCCTCCATGCGCTGTATGACCACTGCTCCCTTACCMAGAGTTTCTTATGAGAAA 6969
Qy 5702 TGACAGCTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761
Db 6970 TGACAGCTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 7029
Qy 5762 CTTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTCATGC 5821
Db 7030 CTTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTCATGC 7089
Qy 5822 ATGCAATACCTCATTTATTGGACAGAAAATGATTAATCTCTAGTTGCTCTGAGGACTTG 5881
Db 7090 ATGCAATACCTCATTTATTGGACAGAAAATGATTAATCTCTAGTTGCTCTGAGGACTTG 7149
Qy 5882 GAGTCACTGCTGTGTGGACTTACTTCAACCAAACTGGTATGTCTCATGAGGCTGGAGTTC 5941
Db 7150 GAGTCACTGCTGTGTGGACTTACTTCAACCAAACTGGTATGTCTCATGAGGCTGGAGTTC 7209
Qy 5942 AAGATCAGGCAAGAGAAAACATGTAAGAAAGTAACTCTCCCACTCACCCGGGTACATG 6001
Db 7210 AAGATCAGGCAAGAGAAAACATGTAAGAAAGTAACTCTCCCACTCACCCGGGTACATG 7269
Qy 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATCAAAACCTCCGTACCC 6061
Db 7270 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATCAAAACCTCCGTACCC 7329
Qy 6062 ATACTCGCTGCTAGGCTATTTAATACCACTCCTCACTGGGCTCCATGAGGTCTCGGCC 6121
Db 7330 ATACTCGCTGCTAGGCTATTTAATAACCACTCCTCACTGGGCTCCATGAGGTCTCGGCC 7389
```



```
QY 6122 AAAACCTACTACTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTCAATCC 6181
Db 7390 AAAACCTACTACTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTCAATCC 7449
QY 6182 CTGTACCTGAACCAATGGAACAACTTCAGCACAGAAATAAAACACACATTCCTGTTTAGTAG 6241
Db 7450 CTGTACCTGAACCAATGGAACAACTTCAGCACAGAAATAAAACACACATTCCTGTTTAGTAG 7509
QY 6242 GACCTCTGTTTCCATSTGGAATAAACCATACCTCAAACTCCTGCTGTTGAAAATTTA 6301
Db 7510 GACCTCTGTTTCCATSTGGAATAAACCATACCTCAAACTCCTGCTGTTGAAAATTTA 7569
QY 6302 GCAATACATACACAAACCAACTCCCAATGCAATCAGGTGGTAACTCCCTCCACACAAA 6361
Db 7570 GCAATACATACACAAACCAACTCCCAATGCAATCAGGTGGTAACTCCCTCCACACAAA 7629
QY 6362 TAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCCTATCGTTGTTGA 6421
Db 7630 TAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCCTATCGTTGTTGA 7689
QY 6422 ATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTAGTGCCCTCCVATGRCATCTACA 6481
Db 7690 ATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTAGTGCCCTCCVATGRCATCTACA 7749
QY 6482 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCTTC 6541
Db 7750 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCTTC 7809
QY 6542 TTCTCTTTTGTATAGGACAGGAGTCTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCA 6601
Db 7810 TTCTCTTTTGTATAGGACAGGAGTCTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCA 7869
QY 6602 CAACCTCTACTAGTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 6661
Db 7870 CAACCTCTACTAGTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 7929
QY 6662 TGCCGACCTCCCTGGTCCACCTTCGAAGTCAACTTAACTCCCTAGCAGCAGTACTCTTC 6721
Db 7930 TGCCGACCTCCCTGGTCCACCTTCGAAGTCAACTTAACTCCCTAGCAGCAGTACTCTTC 7989

RESULT 4
US-09-719-554-3
; Sequence 3, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; FILE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936USOPT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 10499
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-3

Query Match 10.3%; Score 779; DB 3; Length 10499;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 4861
Db 7101 AGGCCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 7160
QY 4862 TAATTCACAGATTCGAGCTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGG 4921
```

```
Db 7161 TAATTCACAGATTCGAGCTTCCCGAGGCTTACAGAGTGCAATAGCCCTGCTTTCCAGG 7220
QY 4922 CCACAGTAACCCAGGAGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 4981
Db 7221 CCACAGTAACCCAGGAGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 7280
QY 4982 GAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA 5041
Db 7281 GAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA 7340
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGGCTGCTGTTGCCCTATAGCCTTAAAAA 5101
Db 7341 AAAAGCAAAACCCAGGAAACCCACCTCACATGGGCTGCTGTTGCCCTATAGCCTTAAAAA 7400
QY 5102 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
Db 7401 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 7460
QY 5162 CTTTCATAACCAATGACCTTGTGCTTGACCCAAAGACAGCAACTTATAGTTCAGACATCAC 5221
Db 7461 CTTTCATAACCAATGACCTTGTGCTTGACCCAAAGACAGCAACTTATAGTTCAGACATCAC 7520
QY 5222 CTCCTTAGCCNAATATCAACAGTTCTTAAAAATTTACAGGAACCTATCCCTGAGAGA 5281
Db 7521 CTCCTTAGCCNAATATCAACAGTTCTTAAAAATTTACAGGAACCTATCCCTGAGAGA 7580
QY 5282 GGGAAAAGAACTATTCCACGCMWGTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC 5341
Db 7581 GGGAAAAGAACTATTCCACGCTTGTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC 7640
QY 5342 CCCATCCCTAGATACATCTCTGGGAAGACCTACCCAGTCACTTTTATATACCCCACTGC 5401
Db 7641 CCCATCCCTAGATACATCTCTGGGAAGACCTACCCAGTCACTTTTATATACCCCACTGC 7700
QY 5402 GGTATAAGTGGCTGGAGTGGAGTCTTGGATATACATACATCTAGTCAAAATCTCTGGATACT 5461
Db 7701 GGTATAAGTGGCTGGAGTGGAGTCTTGGATATACATACATCTAGTCAAAATCTCTGGATACT 7760
QY 5462 GCCAAAGAAACCTTGAAAATCCAGGAGACACCGTAGCTATTCCTGTGMACTCTAGAGGA 5521
Db 7761 GCCAAAGAAACCTTGAAAATCCAGGAGACACCGTAGCTATTCCTGTGMACTCTAGAGGA 7820
QY 5522 TTTGGCCTGCTCTTCAAAACAAACACAGGAGAAAGTAACTAAATCATATAATCCCCCA 5581
Db 7821 TTTGGCCTGCTCTTCAAAACAAACACAGGAGAAAGTAACTAAATCATATAATCCCCCA 7879
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTSTTTTCACTCTCACTG 5641
Db 7880 TGGSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCTCTCTTCACTCTCACTG 7939
QY 5642 CACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTTACCMAGAGTTTCTATGAGAA 5701
Db 7940 CACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTTACCMAGAGTTTCTATGAGAA 7999
QY 5702 TGCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761
Db 8000 TGCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 8059
QY 5762 CTTTCACTGCCCCACACCCCATATGCCCGCAACTGCTATCACTCTGCACTCTTTTGTATGC 5821
Db 8060 CTTTCACTGCCCCACACCCCATATGCCCGCAACTGCTATCACTCTGCACTCTTTTGTATGC 8119
QY 5822 ATGCCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTTG 5881
Db 8120 ATGCCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTTG 8179
QY 5882 GAGTCACTCTGTTGGACTTACTTCCACCAACTGGTATGCTGATGGGGTGGAGTTC 5941
Db 8180 GAGTCACTCTGTTGGACTTACTTCCACCAACTGGTATGCTGATGGGGTGGAGTTC 8239
QY 5942 AAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAACTCCCAACTCACCGSGGTACATG 6001
```

```
Db 8240 AAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCACTCACCCGGGTACATG 8299
Qy 6002 GCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 6061
Db 8300 GCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 8359
Qy 6062 ATACTCGCTGTGTAGAGCTATTATATACACACCTCACTGGGCTCCATGAGGTCTCGGCC 6121
Db 8360 ATACTCGCTGTGTAGAGCTATTATATACACACCTCACTGGGCTCCATGAGGTCTCGGCC 8419
Qy 6122 AAAACCTCTAACTGTGTGGATATGCTCCCTCGAATCTTCARGCCATATGTTTCAATCC 6181
Db 8420 AAAACCTCTAACTGTGTGGATATGCTCCCTCGAATCTTCARGCCATATGTTTCAATCC 8479
Qy 6182 CTGTACTGAAATGGAAACAATTCAGCAGAGAAATAAACACCACTTCCTCCACACAAA 6241
Db 8480 CTGTACTGAAATGGAAACAATTCAGCAGAGAAATAAACACCACTTCCTCCACACAAA 8539
Qy 6242 GACCTCTTGTTCCAATSTGGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTA 6301
Db 8540 GACCTCTTGTTCCAATCTGGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTA 8599
Qy 6302 GCAATACTACATACACAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 6361
Db 8600 GCAATACTACATACACAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 8659
Qy 6362 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGATCTCAGCTTATCGTTTGA 6421
Db 8660 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGATCTCAGCTTATCGTTTGA 8719
Qy 6422 ATGGCTCTTCAGAACTATGTCTCTCACTTCTAGTCCCTCCCTATGACCATCTACA 6481
Db 8720 ATGGCTCTTCAGAACTATGTCTCTCACTTCTAGTCCCTCCCTATGACCATCTACA 8779
Qy 6482 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCATTTC 6541
Db 8780 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCATTTC 8839
Qy 6542 TTCCTTTTGTATAGAGCAGAGTGCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 6601
Db 8840 TTCCTTTTGTATAGAGCAGAGTGCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 8899
Qy 6602 CAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAAACGG 6661
Db 8900 CAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAAACGG 8959
Qy 6662 TCGCGACTCCCTGGTCACTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCCTTC 6721
Db 8960 TCGCGACTCCCTGGTCACTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCCTTC 9019
```

## RESULT 5

```
US-09-949-016-17382
; Sequence 17382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17382
; LENGTH: 13537
; TYPE: DNA
```

```
i ORGANISM: Human
US-09-949-016-17382

Query Match 10.3%; Score 779; DB 3; Length 13537;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

Qy 4802 AGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAAA 4861
Db 8758 AGGCCTTCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAAA 8817
Qy 4862 TAAATTCACAGATTCGGACTTCCCAGAGGCTTTACAGAGTGACAATAAGCCCTGCTTTCCAGG 4921
Db 8818 TAAATTCACAGATTCGGACTTCCCAGAGGCTTTACAGAGTGACAATAAGCCCTGCTTTCCAGG 8877
Qy 4922 CCACAGTAAACCCAGGAGTATCCAGAGGTTAGGTATAGATATACATTACATCGCGCT 4981
Db 8878 CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATAGATATACATTACATCGCGCT 8937
Qy 4982 GAAGGCCACAGTCTCCTCAGGGAAGGTTCGAGAAAATGAATGAACACTCAAGAGCATCTAA 5041
Db 8938 GAAGGCCACAGTCTCCTCAGGGAAGGTTCGAGAAAATGAATGAACACTCAAGAGCATCTAA 8997
Qy 5042 AAAAGCAAAACCCAGGAAAACCCACCTCACATGGCCTGTTCTGTGCTATAGCTTTAAAAA 5101
Db 8998 AAAAGCAAAACCCAGGAAAACCCACCTCACATGGCCTGTTCTGTGCTATAGCTTTAAAAA 9057
Qy 5102 GAAATCTGCAACTTTTCCAAAAGCAGGACTTTAGCCCATACGAAATGCTGTATGGAAGC 5161
Db 9058 GAAATCTGCAACTTTTCCAAAAGCAGGACTTTAGCCCATACGAAATGCTGTATGGAAGC 9117
Qy 5162 CTTTCATAAACCAATGACCTTGTGCTTGACCAAGACAGCCAACTTAGTTGAGACATCAC 5221
Db 9118 CTTTCATAAACCAATGACCTTGTGCTTGACCAAGACAGCCAACTTAGTTGAGACATCAC 9177
Qy 5222 CTCCTTAGCCAAATATCAACAAAGTCTTTAAAAACATTACAAGAACTATCCCTGAGAAGA 5281
Db 9178 CTCCTTAGCCAAATATCAACAAAGTCTTTAAAAACATTACAAGAACTATCCCTGAGAAGA 9237
Qy 5282 GGGAAAAGAACTATTTCCACCCCTTGTGATAGTATAGTCAAGTCCCTTCCTAAATTC 5341
Db 9238 GGGAAAAGAACTATTTCCACCCCTTGTGATAGTATAGTCAAGTCCCTTCCTAAATTC 9297
Qy 5342 CCCATCCCTAGATACATCCTGGGAGGACCTACCCAGTCACTTTTATTTACCCCACTGC 5401
Db 9298 CCCATCCCTAGATACATCCTGGGAGGACCTACCCAGTCACTTTTATTTACCCCACTGC 9357
Qy 5402 GGTAAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATACT 5461
Db 9358 GGTAAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATACT 9417
Qy 5462 GCCAAAGAACTCTGAAAATCCAGGAGCAACGCTAGTATTCCTGTGAACCTCTAGAGA 5521
Db 9418 GCCAAAGAACTCTGAAAATCCAGGAGCAACGCTAGTATTCCTGTGAACCTCTAGAGA 9477
Qy 5522 TTTGGCCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACTAAATCATAAATCCCCA 5581
Db 9478 TTTGGCCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACTAAATCATAAATCCCCA 9536
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCCTSTTTCACTCTCCTG 5641
Db 9537 TGGSCCTCCCTTATCATATTTTCTCTTATCTGTTTACCTCTTCACTCTCCTG 9596
Qy 5642 CACCCCTCCATGCGCTGTATGACCAAGTACTCCCTTACCMAGAGTTTCTATGAGAA 5701
Db 9597 CACCCCTCCATGCGCTGTATGACCAAGTACTCCCTTACCMAGAGTTTCTATGAGAA 9656
Qy 5702 TGCAGCTCCCGGAATATGATGCCCATCGTATAGGAGTCTTTSTAGGAAACCCCA 5761
Db 9657 TGCAGCTCCCGGAATATGATGCCCATCGTATAGGAGTCTTTSTAGGAAACCCCA 9716
Qy 5762 CCTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGTCATGC 5821
```

```
Db 9717 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGGCATGC 9776
Qy 5822 ATCCAAATCACTAATTTGACAGGAAAATGATTAATCTAGTTGCTCTGGAGACTTG 5881
Db 9777 ATGCAATCACTAATTTGACAGGAAAATGATTAATCTAGTTGCTCTGGAGACTTG 9836
Qy 5882 GAGTCACTGCTCTGGACTTACTTCAACCAAACTGCTATGCTGATGGGGTGGAGTTTC 5941
Db 9837 GAGTCACTGCTCTGGACTTACTTCAACCAAACTGCTATGCTGATGGGGTGGAGTTTC 9896
Qy 5942 AGATCAGGCAAGAGAAAACATGTAAAGAACTAATCTCCAACTCACCGGGTACATG 6001
Db 9897 AGATCAGGCAAGAGAAAACATGTAAAGAACTAATCTCCAACTCACCGGGTACATG 9956
Qy 6002 GCACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACATGATAAACCCTCCGTACCC 6061
Db 9957 GCACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACATGATAAACCCTCCGTACCC 10016
Qy 6062 ATACTCGCTGGTAAGCTTATTAATACCAACCTCACTGGGCTCCATGAGTCTCGGCC 6121
Db 10017 ATACTCGCTGGTAAGCTTATTAATACCAACCTCACTGGGCTCCATGAGTCTCGGCC 10076
Qy 6122 AAACCTCTACTAATCTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 6181
Db 10077 AAACCTCTACTAATCTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 10136
Qy 6182 CTGTACTGAACAAATGGAACACTTCAGACAGAAATAAACCACCTCTCGTTTATAGTAG 6241
Db 10137 CTGTACTGAACAAATGGAACACTTCAGACAGAAATAAACCACCTCTCGTTTATAGTAG 10196
Qy 6242 GACCTCTGTTTCCAAATGGAATAAACCATACCTCAACCTCAACCTCACTGTGTAATAATTTA 6301
Db 10197 GACCTCTGTTTCCAAATGGAATAAACCATACCTCAACCTCAACCTCACTGTGTAATAATTTA 10256
Qy 6302 GCAATACTACATACACCAACTCCCAATGCAATCAGTGGGTAACTCTCTCCACACAAA 6361
Db 10257 GCAATACTACATACACCAACTCCCAATGCAATCAGTGGGTAACTCTCTCCACACAAA 10316
Qy 6362 TAGTCTGCTACCTCAGGAATATTTTGTCTGGTACCTCAGCTACCTCAGCTACCTCAGCT 6421
Db 10317 TAGTCTGCTACCTCAGGAATATTTTGTCTGGTACCTCAGCTACCTCAGCTACCTCAGCT 10376
Qy 6422 ATGGCTCTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6481
Db 10377 ATGGCTCTCAGAACTATTTTGTCTGGTACCTCAGCTACCTCAGCTACCTCAGCTACCT 10436
Qy 6482 CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATTC 6541
Db 10437 CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATTC 10496
Qy 6542 TTCTTTTGTATAGAGCAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 6601
Db 10497 TTCTTTTGTATAGAGCAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 10556
Qy 6602 CAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAATGGGGAATGGAACGGG 6661
Db 10557 CAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAATGGGGAATGGAACGGG 10616
Qy 6662 TCGCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGAGTAGTCCCTTC 6721
Db 10617 TCGCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGAGTAGTCCCTTC 10676

RESULT 6
US-09-719-554-22
; Sequence 22, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
```

```
; FILE REFERENCE: 200936USOPCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 22
; LENGTH: 2055
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2055)
; OTHER INFORMATION:
US-09-719-554-22
```

```
Query Match 9.6%; Score 731; DB 3; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5685 AGAGTTTCTATGAGAAATGCAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCT 5744
Db 495 AGAGTTTCTATGAGAAATGCAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCT 554
Qy 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 5804
Db 555 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 614
Qy 5805 TGCCACTCTTTGCATGCAATGCAATCACTATTTGGACAGGAAATGATTAATCTCTAG 5864
Db 615 TGCCACTCTTTGCATGCAATGCAATCACTATTTGGACAGGAAATGATTAATCTCTAG 674
Qy 5865 TTGCTCTGGAGACTTTGGAGTCACTGCTGTTGGACTTACTTACCCAACTGGTATGTC 5924
Db 675 TTGCTCTGGAGACTTTGGAGTCACTGCTGTTGGACTTACTTACCCAACTGGTATGTC 734
Qy 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACAATGTAAAGAAAGTAACTCTCCCA 5984
Db 735 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACAATGTAAAGAAAGTAACTCTCCCA 794
Qy 5985 ACTCACCGGGTACATGGCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTACA 6044
Db 795 ACTCACCGGGTACATGGCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTACA 854
Qy 6045 TGAACCCCTCGTACCCATCTCGCTCTGTAAGCCTTATTAATACCCCTCAGCTGGGT 6104
Db 855 TGAACCCCTCGTACCCATCTCGCTCTGTAAGCCTTATTAATACCCCTCAGCTGGGT 914
Qy 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCTCCCTCGAATTCAR 6164
Db 915 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCTCCCTCGAATTCAR 974
Qy 6165 GCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTTCAGACAGAAATAAACAC 6224
Db 975 GCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTTCAGACAGAAATAAACAC 1034
Qy 6225 CACTTCCGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACT 6284
Db 1035 CACTTCCGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACT 1094
Qy 6285 CACTTGTGTAAATTTTAGCAATACTACATACAAACCAACTCCCAATGATCAGGTGGGT 6344
Db 1095 CACTTGTGTAAATTTTAGCAATACTACATACAAACCAACTCCCAATGATCAGGTGGGT 1154
Qy 6345 AACTCTCCCAACAATAATAGTCTGCTACCTCAGGAATAATTTTTGTCTGTGGTACCTC 6404
Db 1155 AACTCTCCCAACAATAATAGTCTGCTACCTCAGGAATAATTTTTGTCTGTGGTACCTC 1214
Qy 6405 AGCCTATCGTTGTTGAATGAGTCTTTCAGAACTATGCTCTCTCATTTCTTAGTGCC 6464
Db 1215 AGCCTATCGTTGTTGAATGAGTCTTTCAGAACTATGCTCTCTCATTTCTTAGTGCC 1274
```

```

QY 6465 CCCATGTCCTACCTGAAACAAATTTATACAGTTATGTCATATCTAAGCCCGCAA 6524
Db 1275 CCCTATGACCATCTACATCTGAACAGATTTATACAGTTATGTCATATCTAAGCCCGCAA 1334
QY 6525 CAAAGAGTACCCATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGACCTAGGTAC 6584
Db 1335 CAAAGAGTACCCATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGACCTAGGTAC 1394
QY 6585 TGGCATTCGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1395 TGGCATTCGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1454
QY 6645 TGGGACATGGAACGGGTGCGGACTCCTCTGTGTACCTTGGCAAGATCAACTTAACTCCCT 6704
Db 1455 TGGGACATGGAACGGGTGCGGACTCCTCTGTGTACCTTGGCAAGATCAACTTAACTCCCT 1514
QY 6705 AGCAGCAGTAGTCTTTC 6721
Db 1515 AGCAGCAGTAGTCTTTC 1531

RESULT 7
US-09-949-016-5640
; Sequence 5640, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5640
; LENGTH: 2763
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5640

Query Match 9.6%; Score 731; DB 3; Length 2763;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGAAATGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 867 AGAGTTTCTATGGAGAAATGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 926
QY 5745 TTSTAAGGAAACCCCATCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804
Db 927 TTSTAAGGAAACCCCATCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 986
QY 5805 TGCCACTCTTTGCATGCATGCAAACTACTATTATGGACAGGAAATGATTAATCTCTAG 5864
Db 987 TGCCACTCTTTGCATGCATGCAAACTACTATTATGGACAGGAAATGATTAATCTCTAG 1046
QY 5865 TTGTCTTGAGGACHTGGAGTCACTGTCTGTGGACTTACTTCCACCAAACTGGTATGTC 5924
Db 1047 TTGTCTTGAGGACHTGGAGTCACTGTCTGTGGACTTACTTCCACCAAACTGGTATGTC 1106
QY 5925 TGATGGGGTGGAGTTCAAGATCAGCAGGAGAGAAAACATGTAAGAAGTAATCTCCCA 5984
Db 1107 TGATGGGGTGGAGTTCAAGATCAGCAGGAGAGAAAACATGTAAGAAGTAATCTCCCA 1166
QY 5985 ACTCACCCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 6044
Db 1167 ACTCACCCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 1226

```

```

QY 6045 TGAACCCCTCGTACCCATACCTCGCTGGTAAGCCTATTTAATACCACCCCTCAGTGGCT 6104
Db 1227 TGAACCCCTCGTACCCATACCTCGCTGGTAAGCCTATTTAATACCACCCCTCAGTGGCT 1286
QY 6105 CCATGAGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAR 6164
Db 1287 CCATGAGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAG 1346
QY 6165 GCCATATGTTCAATTCCTGTACTGAAACAATGGAACAACCTTCAGCAAGAAAATAAACAC 6224
Db 1347 GCCATATGTTCAATTCCTGTACTGAAACAATGGAACAACCTTCAGCAAGAAAATAAACAC 1406
QY 6225 CACTTCGGTTTTAGTAGGACCTCTGTTTCCAATSTGGAATTAACCCATACCTCAAACT 6284
Db 1407 CACTTCGGTTTTAGTAGGACCTCTGTTTCCAATSTGGAATTAACCCATACCTCAAACT 1466
QY 6285 CACCTGTGTAATTTTAGCAATACTACATACAAACCAACTCCCAATGCATCAGTGGGT 6344
Db 1467 CACCTGTGTAATTTTAGCAATACTACATACAAACCAACTCCCAATGCATCAGTGGGT 1526
QY 6345 AACTCTCCCAACAATAAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404
Db 1527 AACTCTCCCAACAATAAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1586
QY 6405 AGCCTATCGTTGTTTGAATGCTCTTCAGATCTATGCTCTCTCTCACTTCTTAGTGCC 6464
Db 1587 AGCCTATCGTTGTTTGAATGCTCTTCAGATCTATGCTCTCTCTCACTTCTTAGTGCC 1646
QY 6465 CCCATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAA 6524
Db 1647 CCCATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAA 1706
QY 6525 CAAAAGAGTACCCATCTTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 6584
Db 1707 CAAAAGAGTACCCATCTTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 1766
QY 6585 TGCCATTTGGCGGTATCAACCTCTACTCAGTTCTTACTTACAACTATCTCAAGAACTAAA 6644
Db 1767 TGCCATTTGGCGGTATCAACCTCTACTCAGTTCTTACTTACAACTATCTCAAGAACTAAA 1826
QY 6645 TGGGACATGGAACGGGTGCGGACTCCTCTGTGTACCTTGGCAAGATCAACTTAACTCCCT 6704
Db 1827 TGGGACATGGAACGGGTGCGGACTCCTCTGTGTACCTTGGCAAGATCAACTTAACTCCCT 1886
QY 6705 AGCAGCAGTAGTCTTTC 6721
Db 1887 AGCAGCAGTAGTCTTTC 1903

RESULT 8
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA

```

; ORGANISM: Homo sapiens  
US-09-175-928-3

Query Match 9.0%; Score 680; DB 3; Length 2946;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 5685 AGATTTCTATGAGAAATGAGCGTCCCGGAATATTGATGCCCATCTGATAGGAGTCT 5744
DB 1032 AGAGTTTCTATGAGAAATGAGCGTCCCGGAATATTGATGCCCATCTGATAGGAGTCT 1091
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804
DB 1092 TTCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 1151
QY 5805 TGCACACTTTTGATGATGCAATATCTATTTATGGACAGGAAAAATGATTATCTAG 5864
DB 1152 TGCACACTTTTGATGATGCAATATCTATTTATGGACAGGAAAAATGATTATCTAG 1211
QY 5865 TTGCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGTC 5924
DB 1212 TTGCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGTC 1271
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTCCCA 5984
DB 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTCCCA 1331
QY 5985 ACTCACSSGGTACATGGCACTTACCCCTCAAGGACTAGATCTCTCAAAACTACA 6044
DB 1332 ACTCACSSGGTACATGGCACTTACCCCTCAAGGACTAGATCTCTCAAAACTACA 1391
QY 6045 TGAACCCCTCCGACCATACTCGCTGTAAGCTTATTAATACACCCCTCACTGGCT 6104
DB 1392 TGAACCCCTCCGACCATACTCGCTGTAAGCTTATTAATACACCCCTCACTGGCT 1451
QY 6105 CCATGAGTCTCGGCCCAAAACCTACTACTGTTGATATGCTCCCTCGTGAATTCAR 6164
DB 1452 CCATGAGTCTCGGCCCAAAACCTACTACTGTTGATATGCTCCCTCGTGAATTCAG 1511
QY 6165 GCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCACAGAAATAACAC 6224
DB 1512 GCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCACAGAAATAACAC 1571
QY 6225 CACTTCGGTTTATGAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT 6284
DB 1572 CACTTCGGTTTATGAGGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACT 1631
QY 6285 CACTGTGTAAATTTAGCAATACTACATACAAACCACTCCCAATGCATCAGTGGGT 6344
DB 1632 CACTGTGTAAATTTAGCAATACTACATACAAACCACTCCCAATGCATCAGTGGGT 1691
QY 6345 AACTCTCCACACAATAATGCTGCTTACCCCTCAGGAATATTTTGTCTGTGCTACCTC 6404
DB 1692 AACTCTCCACACAATAATGCTGCTTACCCCTCAGGAATATTTTGTCTGTGCTACCTC 1751
QY 6405 AGCCTATGTTTGAATGGCTCTTCAGAACTATGTTGCTTCCCTCTCAATCTTAGTGCC 6464
DB 1752 AGCCTATGTTTGAATGGCTCTTCAGAACTATGTTGCTTCCCTCTCAATCTTAGTGCC 1811
QY 6465 CCCVATGCCATCTACAGAAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 6524
DB 1812 CCCVATGCCATCTACAGAAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 1871
QY 6525 CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACAGGAGTCTAGGTGCATCTAGTAC 6584
DB 1872 CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACAGGAGTCTAGGTGCATCTAGTAC 1931
QY 6585 TGGCATTTGGCGGTATCACAACTCTACTAGTTCTACTACAACTATCTCAAGAACTAAA 6644
DB 1932 TGGCATTTGGCGGTATCACAACTCTACTAGTTCTACTACAACTATCTCAAGAACTAAA 1991
QY 6645 TGGGACATGGAACGGGTGCGGACCTCCCTGGTCACTTGGCAAGATCAACTTAATCCCT 6704
```

```
DB 1992 TGGGACATGGAACGGGTGCGGACCTCCCTGGTCACTTGGCAAGATCAACTTAATCCCT 2051
QY 6705 AGCAGCAGTAGTCTCTTC 6721
DB 2052 AGCAGCAGTAGTCTCTTC 2068
```

## RESULT 9

US-09-719-554-2  
; Sequence 2, Application US/09719554  
; Patent No. 6919438  
; GENERAL INFORMATION:  
; APPLICANT: ALLIEL, Patrick  
; APPLICANT: PERIN, Jean-Pierre  
; APPLICANT: RIEGER, Francois  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH H  
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS  
; FILE REFERENCE: 200916US0PCT  
; CURRENT APPLICATION NUMBER: US/09/719,554  
; CURRENT FILING DATE: 2001-12-26  
; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
; PRIOR FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1326  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-719-554-2

Query Match 3.5%; Score 267; DB 3; Length 1326;  
Best Local Similarity 99.4%; Pred. No. 3.8e-124;  
Matches 487; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
QY 1507 GCGAAGGGAGAAAATCTACTGCTTCTGGAGAGACTAAAGGAGGCATTGAGGAGCG 1566
DB 779 GCGAAGGGAGAAAATCTACTGCTTCTGGAGAGACTAAAGGAGGCATTGAGGAGCG 838
QY 1567 TGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAACTCTTAAAGCGTAAGTTTATCAC 1626
DB 839 TGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAACTCTTAAAGCGTAAGTTTATCAC 898
QY 1627 TGATCAGCTGCGACAGATTAG-AAAAAATTCAAAGTCTGCCGTAGGCCCGGAGCAAAA 1685
DB 899 TCAGTCAGCTGCGACAGATTAGAAAAAACTTCAAAGTCTGCCGTAGGCCCGGAGCAAAA 958
QY 1686 CTTAGAAACCTATTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA 1745
DB 959 CTTAGAAACCTATTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA 1018
QY 1746 GCGGAACAGGACAAACGGGATTAAAAAAGGCCACCCCTTTAGTCATGACCTCAGGC 1805
DB 1019 GCGGAACAGGACAAACGGGATTAAAAAAGGCCACCCCTTTAGTCATGACCTCAGGC 1078
QY 1806 AAGTGAACCTTTGAGGCTCTGAAAAGGAAAAGCTGGGCAAAATTGAAATGCCCTAATAGGG 1865
DB 1079 AAGTGAACCTTTGAGGCTCTGAAAAGGAAAAGCTGGGCAAAATTGAAATGCCCTAATAGGG 1138
QY 1866 CTTGCTTCCAGTCCGCTCTACAAGGACACTTTAAAAAAGATTTGTCCAAGTAGAAGTAAGC 1925
DB 1139 CTTGCTTCCAGTCCGCTCTACAAGGACACTTTAAAAAAGATTTGTCCAAGTAGAAGTAAGC 1198
QY 1926 CGCCCTCTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAGGCCCACTGCCCGAGG 1985
DB 1199 CGCCCTCTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAGGCCCACTGCCCGAGG 1258
QY 1986 GACAAAGGTC 1995
DB 1259 GACAAAGGTC 1268
```

## RESULT 10

US-09-719-554-20

```
; Sequence 20, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIED, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-20

Query Match      2.8%; Score 211; DB 3; Length 711;
Best Local Similarity 99.6%; Pred. No. 86-96;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60
Db 209 CAACATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 268

QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACTCTATTAAATCTTGCAACTGCA 120
Db 269 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACTCTATTAAATCTTGCAACTGCA 328

QY 121 CTCCTTGTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACCTGC 180
Db 329 CTCCTTGTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACCTGC 388

QY 181 TGTTTGCCACACCCGACANACCTCCGCTGACTCCCATCCCTCTGGATCTGCGAGGTGC 240
Db 389 TGTTTGCCACACCCGACAGACCTGCGCTGACTCCCATCCCTCTGGATCTGCGAGGTGC 448

QY 241 CGCTGTGCTCCTGATCCAGCGA 262
Db 449 CGCTGTGCTCCTGATCCAGCGA 470

RESULT 11
US-08-686-878A-48
; Sequence 48, Application US/08686878A
; Patent No. 5708157
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Bridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn version 3.1
; CURRENT APPLICATION NUMBER: US/08/721,489
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,878A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-686-878A-48

Query Match      2.5%; Score 190; DB 2; Length 542;
Best Local Similarity 100.0%; Pred. No. 3.3e-85;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5390 TACCCCAACTCGGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAA 5449
Db 152 TACCCCAACTCGGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAA 211

QY 5450 ATCTCGGATATGCAAGGAACCTGAAATCCAGGAGACAACTAGCTATTCCTGTGA 5509
Db 212 ATCTCGGATATGCAAGGAACCTGAAATCCAGGAGACAACTAGCTATTCCTGTGA 271

QY 5510 ACCTCTAGAGATTGGCCCTGCTCTTCAAAACAACACAGGAGAAAGTAACTAAATC 5569
Db 272 ACCTCTAGAGATTGGCCCTGCTCTTCAAAACAACACAGGAGAAAGTAACTAAATC 331

QY 5570 ATAAATCCCC 5579
Db 332 ATAAATCCCC 341

RESULT 12
US-08-721-489-2
; Sequence 2, Application US/08721489
; Patent No. 5786465
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Bridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,489
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-721-489-2

Query Match 2.5%; Score 190; DB 2; Length 542;
Best Local Similarity 100.0%; Pred. No. 3.3e-85;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5390 TACCCCAACTGCGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAA 5449
DB 152 TACCCCAACTGCGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAA 211
QY 5450 ATCCTGATATCTCCAAAGGAACCTGAAATCCAGGAGACAACGCTAGCTATTTCCTGTGA 5509
DB 212 ATCCTGATATCTCCAAAGGAACCTGAAATCCAGGAGACAACGCTAGCTATTTCCTGTGA 271
QY 5510 ACCTCTAGAGGATTGGCGCTCTTCAAAACCAACAGGAGGAAGTAATAAATC 5569
DB 272 ACCTCTAGAGGATTGGCGCTCTTCAAAACCAACAGGAGGAAGTAATAAATC 331
QY 5570 ATAAATCCCC 5579
DB 332 ATAAATCCCC 341

RESULT 13
US-09-719-554-28
; Sequence 28, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936USOPT
; CURRENT APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 28
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1080)
; OTHER INFORMATION:
; US-09-719-554-28

Query Match 2.4%; Score 183; DB 3; Length 1080;
Best Local Similarity 99.4%; Pred. No. 1.2e-81;
Matches 353; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1507 GGAAGGGAGAAATCTCTACTGCGCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGCG 1566
DB 726 GGAAGGGAGAAATCTCTACTGCGCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGCG 785
QY 1567 TGCCTCTCTGACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTATAC 1626
DB 786 TGCCTCTCTGACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTATAC 845
```

```
QY 1627 TCAGTCAGCTGCAGACATTAG-AAAAAAGCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAA 1685
DB 846 TCAGTCAGCTGCAGACATTAGAAAAAAGCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAAA 905
QY 1686 CTTAGAAACCCCTATTCAACTTGGCAACCTCGGTTTATTAATAGAGATCAGAGGAGCA 1745
DB 906 CTTAGAAACCCCTATTGAACCTTGGCAACCTCGGTTTATTAATAGAGATCAGAGGAGCA 965
QY 1746 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGC 1805
DB 966 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGC 1025
QY 1806 AAGTGAAGCTTGGAGCTCTGAAAAAGGGAAGCTGGGCAAAATTGAATGCCTAA 1860
DB 1026 AAGTGAAGCTTGGAGCTCTGAAAAAGGGAAGCTGGGCAAAATTGAATGCCTAA 1080

RESULT 14
US-09-513-999C-21394
; Sequence 21394, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pn
; SEQ ID NO 21394
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 102
; OTHER INFORMATION: k=g or t
; US-09-513-999C-21394
```

```
Query Match 2.0%; Score 149; DB 3; Length 443;
Best Local Similarity 99.2%; Pred. No. 1.8e-64;
Matches 249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 12 ATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCCCTTCCCTTTGT 71
DB 3 ATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCCCTTCCCTTTGT 62
QY 72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTGCAACTGCACTCTTCTGGTC 131
DB 63 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTGCAACTGCACTCTTCTGGTC 122
QY 132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGCTGTTGCGACC 191
DB 123 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGCTGTTGCGACC 182
QY 192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCTCTGAGGTGTCCGCTGTGCTCC 251
DB 183 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCTCTGAGGTGTCCGCTGTGCTCC 242
QY 252 TCATCCAGCA 262
DB 243 TCATCCAGCA 253
```

```
RESULT 15
US-09-719-554-21
; Sequence 21, Application US/09719554
```



```

; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUMAN HLA CLASS II ANTIGEN-RELATED RETROVIRAL MOTIFS
; FILE REFERENCE: 200936USOPCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-719-554-21

```

	Query Match	1.2%	Score 89	DB 3	Length 711	
	Best Local Similarity	100.0%	Pred. No. 4.1e-34			
	Matches 89	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
Qy	7488	TTGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTCTGTTATCGGAGCTCTGTTTT	7547			
Db	235	TTGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTCTGTTATCGGAGCTCTGTTTT	294			
Qy	7548	CATGCTATTTCACTCTATTAAACTTGCA	7576			
Db	295	CATGCTATTTCACTCTATTAAACTTGCA	323			

RESULT 16  
 US-09-719-554-66  
 ; Sequence 66, Application US/09719554  
 ; Patent No. 6919438  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALLIEL, Patrick  
 ; APPLICANT: PERIN, Jean-Pierre  
 ; APPLICANT: RIEGER, Francois  
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH H  
 ; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS  
 ; FILE REFERENCE: 200936USOPCT  
 ; CURRENT APPLICATION NUMBER: US/09/719,554  
 ; CURRENT FILING DATE: 2001-12-26  
 ; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
 ; PRIOR FILING DATE: 1999-06-25  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 66  
 ; LENGTH: 46340  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-719-554-66

	Query Match	1.0%;	Score 76;	DB 3;	Length 46340;
	Best Local Similarity	100.0%;	Pred. No. 1.7e-27;		
	Matches	76;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	6068	GCCTGGTAA	GCCCTATT	TATATACAC	CCTCAGCTGGGCTCCATGAGGTCTCGGCCCAAACC 6127
Db	32342	GCCTGGTAA	GCCCTATT	TATATACAC	CCTCAGCTGGGCTCCATGAGGTCTCGGCCCAAACC 32401
Qy	6128	CTACTAACT	GTTCGAT	6143	
Db	32402	CTACTAACT	GTTCGAT	32417	

RESULT 17  
US-08-979-847B-105  
; Sequence 105, Application US/08979847B  
; Patent No. 6582703

```

GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
           BESEME, FREDERIC
           BEDIN, FREDERIC
           PARANHOS-BACCALA, GLAUCIA
           KOMURIAN-PRADEL, FLORENCE
           JOLIVET-REYNAUD, COLETTE
           MANDRAND, BERNARD
           GARSON, JEREMY
           TUKE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
                    ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC
                    THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESS: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-No. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-08-979-847B-105

```

	Query Match	1.0%;	Score 73;	DB 3;	Length 1481;
	Best Local Similarity	99.2%;	Pred. No. 5.2e-26;		
	Matches 123;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	6187	CTGTGAACAATGGGAACAACCTTCAGCACAGAAAATAAACACCACTTCCGCTTTTGTAGTAGGACCT	6246		
Db	607	CTGTGAACAATGGGAACAACCTTCAGCACAGAAAATAAACACCACTTCCGCTTTTGTAGTAGGACCT	666		
QY	6247	CTTGTGTTTCCAATSTGGAAATAACCCATACCTCAAACTCACCCTGTGTAAAAATTTAGCAAT	6306		
Db	667	CTTGTGTTTCCAATCTGGAAATAACCCATACCTCAAACTCACCCTGTGTAAAAATTTAGCAAT	726		
QY	6307	ACTA 6310			
Db	727	ACTA 730			

```

RESULT 18
US-09-949-016-13002/c
; Sequence 13002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THERE
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL001307

```

; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 13002  
; LENGTH: 168394  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(168394)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13002

Query Match 0.8%; Score 64; DB 3; Length 168394;  
Best Local Similarity 100.0%; Pred. No. 2e-21;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1999 TGAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGTGCTGGGCAAGC 2058  
|||||  
Db 28029 TGAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGTGCTGGGCAAGC 27970  
|||||  
QY 2059 GCCA 2062  
|||||  
Db 27969 GCCA 27966

## RESULT 19

US-09-719-554-5/c  
; Sequence 5, Application US/09719554  
; Patent No. 6919438  
; GENERAL INFORMATION:  
; APPLICANT: ALLIEL, Patrick  
; APPLICANT: PERIN, Jean-Pierre  
; APPLICANT: RIEGER, Francois  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU  
; FILE REFERENCE: 200936US0PCT  
; CURRENT APPLICATION NUMBER: US/09/719,554  
; CURRENT FILING DATE: 2001-12-26  
; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
; PRIOR FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1799  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-719-554-5

Query Match 0.8%; Score 63; DB 3; Length 1799;  
Best Local Similarity 100.0%; Pred. No. 1e-21;  
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATCTTAGTGCC 6464  
Db 938 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATCTTAGTGCC 879  
QY 6465 CCC 6467  
|||  
Db 878 CCC 876

## RESULT 20

US-09-949-016-15858/c  
; Sequence 15858, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 15858  
; LENGTH: 145320  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15858

Query Match 0.8%; Score 63; DB 3; Length 145320;  
Best Local Similarity 100.0%; Pred. No. 6.5e-21;  
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATCTTAGTGCC 6464  
|||||  
Db 93202 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATCTTAGTGCC 93143  
|||||  
QY 6465 CCC 6467  
|||  
Db 93142 CCC 93140

## RESULT 21

US-08-979-847B-112  
; Sequence 112, Application US/08979847B  
; Patent No. 6582703  
; GENERAL INFORMATION:  
; APPLICANT: PERRON, HERVE  
; BESEME, FREDERIC  
; BEDIN, FREDERIC  
; PARANHOS-BACCALA, GLAUCIA  
; KOMURIAN-PRADEL, FLORENCE  
; JOLIVET-REYNAUD, COLETTE  
; MANDRAND, BERNARD  
; GARSON, JEREMY  
; TUBE, PHILIP  
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY  
; THERAPEUTIC PURPOSES  
  
NUMBER OF SEQUENCES: 210  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. BOX 19928  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,847B  
FILING DATE: 26-No. 6582703-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BERRIDGE, WILLIAM P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 39046A  
TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 112:
US-08-979-847B-112

Query Match          0.8%; Score 58; DB 3; Length 758;
Best Local Similarity 100.0%; Pred. No. 2e-18;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5199 GCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCCTTAAACAT 5256
Db 645 GCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCCTTAAACAT 702

RESULT 22
US-09-949-016-17417/c
; Sequence 17417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17417
; LENGTH: 7772
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(7772)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17417

Query Match          0.7%; Score 56; DB 3; Length 7772;
Best Local Similarity 100.0%; Pred. No. 2.2e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5205 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCCTTAAACATTACA 5260
Db 16566 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCCTTAAACATTACA 16511

RESULT 23
US-09-949-016-12249/c
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 7797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(7797)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

Query Match          0.7%; Score 56; DB 3; Length 7797;
Best Local Similarity 100.0%; Pred. No. 2.2e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5205 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCCTTAAACATTACA 5260
Db 16570 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCCTTAAACATTACA 16515

RESULT 24
US-10-104-047-1148
; Sequence 1148, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1148
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1148

Query Match          0.7%; Score 55; DB 3; Length 2349;
Best Local Similarity 100.0%; Pred. No. 6.7e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7232 GGTTCCTCTGTGAGATGGGGACTGAGACAGGACTAGCTGGATTTCCTAGGC 7286
Db 1626 GGTTCCTCTGTGAGATGGGGACTGAGACAGGACTAGCTGGATTTCCTAGGC 1680

RESULT 25
US-09-949-016-13590/c
; Sequence 13590, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13590
```

; LENGTH: 149971  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(149971)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13590

Query Match 0.7%; Score 55; DB 3; Length 149971;  
Best Local Similarity 100.0%; Pred. No. 7.3e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1999 TGAGTCAGAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGG 2053  
Db 19748 TGAGTCAGAGCCACTAACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGG 19694

## RESULT 26

US-09-979-847B-117  
; Sequence 117, Application US/08979847B  
; Patent No. 6582703  
; GENERAL INFORMATION:  
; APPLICANT: PERRON, HERVE  
; BESEME, FREDERIC  
; BEDIN, FREDERIC  
; PARANHOS-BACCALA, GLAUCIA  
; KOMURIAN-PRADEL, FLORENCE  
; JOLIVET-REYNAUD, COLETTE  
; MANDRAND, BERNARD  
; GARSON, JEREMY  
; TUREK, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL  
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. BOX 19928  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22320

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,847B  
FILING DATE: 26-No. 6582703-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BERRIDGE, WILLIAM P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 39046A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6400  
TELEFAX: 703-836-2787

## INFORMATION FOR SEQ ID NO: 117:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1511 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 117:  
US-08-979-847B-117

Query Match 0.7%; Score 54; DB 3; Length 1511;  
Best Local Similarity 100.0%; Pred. No. 2.1e-16;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1615 TAAGTTTATCACTCACTGAGTCGAGACATTAGAAAAAACTTCAAAAGTCTGCC 1668  
Db 1003 TAAGTTTATCACTCACTGAGTCGAGACATTAGAAAAAACTTCAAAAGTCTGCC 1056

## RESULT 27

US-09-403-343B-32  
; Sequence 32, Application US/09403343B  
; Patent No. 6555091  
; GENERAL INFORMATION:  
; APPLICANT: JOLIVET-REYNAUD, COLETTE  
; APPLICANT: PERRON, HERVE  
; APPLICANT: MANDRAND, BERNARD  
; TITLE OF INVENTION: POLYPEPTIDE CAPABLE OF REACTING WITH ANTIBODIES OF  
; FILE REFERENCE: 104574  
; CURRENT APPLICATION NUMBER: US/09/403,343B  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: FR/97/05679  
; PRIOR FILING DATE: 1997-04-29  
; PRIOR APPLICATION NUMBER: FR/97/16870  
; PRIOR FILING DATE: 1997-12-31  
; PRIOR APPLICATION NUMBER: PCT/FR98/00870  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 32  
; LENGTH: 286  
; TYPE: DNA  
; ORGANISM: MSRV  
; US-09-403-343B-32

Query Match 0.7%; Score 53; DB 3; Length 286;

Best Local Similarity 100.0%; Pred. No. 6.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1513 GGAAGAAAATCTACTGCTTCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGC 1565  
Db 12 GGAAGAAAATCTACTGCTTCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGC 64

## RESULT 28

US-08-691-563C-88  
; Sequence 88, Application US/08691563C  
; Patent No. 6001987  
; GENERAL INFORMATION:

APPLICANT: Herve PERRON  
APPLICANT: Frederic BESEME  
APPLICANT: Frederic BEDIN  
APPLICANT: Glaucia PARANHOS-BACCALA  
APPLICANT: Florence KOMURIAN-PRADEL  
APPLICANT: Colette JOLIVET  
APPLICANT: Bernard MANDRAND  
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLAC  
NUMBER OF SEQUENCES: 92  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Oliff & Berridge  
STREET: 700 South Washington Street, Suite 300  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22314

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/691,563C  
FILING DATE: 02-AUG-1996  
ATTORNEY/AGENT INFORMATION:

```
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 693 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-691-563C-88

Query Match      0.7%; Score 53; DB 3; Length 693;
Best Local Similarity 100.0%; Pred. No. 6.7e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1513 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565
Db 577 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 629

RESULT 29
US-09-374-766-88
; Sequence 88, Application US/09374766
; Patent No. 6579526
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Frederic BESEME
; APPLICANT: Frederic BEDIN
; APPLICANT: Glaucaia PARANHOS-BACCALA
; APPLICANT: Florence KOMURIAN-PRADEL
; APPLICANT: Colette JOLIVET
; APPLICANT: Bernard MANDRAND
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/374,766
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,563
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 693 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-374-766-88
```

```
Query Match      0.7%; Score 53; DB 3; Length 693;
Best Local Similarity 100.0%; Pred. No. 6.7e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1513 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565
Db 577 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 629

RESULT 30
US-08-979-847B-82
; Sequence 82, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKER, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 693 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 82:
; US-08-979-847B-82

Query Match      0.7%; Score 53; DB 3; Length 693;
Best Local Similarity 100.0%; Pred. No. 6.7e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1513 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565
Db 577 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 629

Search completed: August 2, 2006, 13:30:03
Job time : 1256 secs
```

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:55 ; Search time 21273 Seconds  
(without alignments)  
19930.446 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7582  
Sequence: 1 caacaatcgggataataaac.....tattaaatcttgcactgcr 7582

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96469239

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 90 summaries

```
Database : EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Query %			DB	ID	Description			
	Score	Match	Length						
1	731	9.6	2500	6	CR622175	CR622175	full-length		
2	731	9.6	2716	6	CR613169	CR613169	full-length		
3	731	9.6	2748	6	CR605851	CR605851	full-length		
4	731	9.6	2748	6	CR625046	CR625046	full-length		
5	731	9.6	2749	6	CR617248	CR617248	full-length		
C	6	525	935	4	BX331741	BX331741	full-length		
	7	446	6.9	828	4	CR6367907	CR6367907		
	8	442	5.8	903	4	EX408752	EX408752		
	9	437	5.8	883	4	EX408752	EX408752		
C	10	431	5.7	872	4	EX408733	EX408733		
	11	411	5.4	931	4	EX4326647	EX4326647		
	12	403	5.3	792	9	CR389656	CR389656		
	13	379	5.0	564	9	DA8300320	DA8300320		
	14	377	5.0	924	4	CR436641	CR436641		
	15	374	4.9	538	9	DA838045	DA838045		
	16	373	4.9	721	1	A0138405	A0138405		
	17	373	4.9	761	9	DA832031	DA832031		
	18	372	4.9	540	9	DA849357	DA849357		
	19	370	4.9	819	9	DA830049	DA830049		

## ALIGNMENTS

```
RESULT 1
CR622175
LOCUS
DEFINITION
full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized
of Homo sapiens (human) .
ACCESSION
CR622175
VERSION
CR622175.1 GI:50502982
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2500)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 2500)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..2500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 9.6%; Score 731; DB 6; Length 2500;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 5685 AGAGTTTCTATGGAGATCGAGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 634 AGAGTTTCTATGGAGATCGAGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 693
QY 5745 TTSTAGGGAACCCACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804
Db 694 TTCTAAGGGAACCCACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 753
QY 5805 TGCCACTCTTTGATCATGCAATCACTATTTATGGACAGGAAATGATTAATCTCTAG 5864
Db 754 TGCCACTCTTTGATCATGCAATCACTATTTATGGACAGGAAATGATTAATCTCTAG 813
QY 5865 TTGTCTCGAGGACTGGAGTCACTGCTGTGGACTTACTTCAACCACCTGGTATGTC 5924
Db 814 TTGTCTCGAGGACTGGAGTCACTGCTGTGGACTTACTTCAACCACCTGGTATGTC 873
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAAGTAAATCTCCCA 5984
Db 874 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAAGTAAATCTCCCA 933
QY 5985 ACTACCSGGGTACATGGGACCTCTAGCCCTTCAAGAGCTAGATCTCTCAAAACTACA 6044
Db 934 ACTACCSGGGTACATGGGACCTCTAGCCCTTCAAGAGCTAGATCTCTCAAAACTACA 993
QY 6045 TGAACCCCTCGGTACCCATACCTCGCTGGTAAAGCCTATTATTAACACCCCTCACTGGGCT 6104
```

```
Db 994 TGAACCCCTCGGTACCCATACCTCGCTGGTAAAGCCTATTATTAACACCCCTCACTGGGCT 1053
QY 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATGCTCCGCCCTGAACCTTCA 6164
Db 1054 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATGCTCCGCCCTGAACCTTCA 1113
QY 6165 GCCATATGTTTCAATCCCTGTACCTGAAACAACTTTCAGCAAGAAATAAACAAC 6224
Db 1114 GCCATATGTTTCAATCCCTGTACCTGAAACAACTTTCAGCAAGAAATAAACAAC 1173
QY 6225 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACT 6284
Db 1174 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACT 1233
QY 6285 CACTGTGTAAATTTTAGCAATATACATACACAACTCCCAATGATCAGTGGGT 6344
Db 1234 CACTGTGTAAATTTTAGCAATATACATACACAACTCCCAATGATCAGTGGGT 1293
QY 6345 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTGCTGTGGTACCTC 6404
Db 1294 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTGCTGTGGTACCTC 1353
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGC 6464
Db 1354 AGCCTATCGTTGTTGAATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGC 1413
QY 6465 CCCATGATGCAATCTACACTGAAACAAAGATTTATACAGTTATGTATATCTAAGCCCGCAA 6524
Db 1414 CCCTATGACCATCTACACTGAAACAAAGATTTATACAGTTATGTATATCTAAGCCCGCAA 1473
QY 6525 CAAAGAGTACCCATCTTCTTCTTGTATAGGACGAGTGTAGTGCATAGGTAC 6584
Db 1474 CAAAGAGTACCCATCTTCTTCTTGTATAGGACGAGTGTAGTGCATAGGTAC 1533
QY 6585 TGGCATTTGGCGGTATCACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 6644
Db 1534 TGGCATTTGGCGGTATCACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 1593
QY 6645 TGGGACATGAAACGGGTGCGGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6704
Db 1594 TGGGACATGAAACGGGTGCGGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653
QY 6705 AGCAGCAGTGTCTTTC 6721
Db 1654 AGCAGCAGTGTCTTTC 1670
RESULT 2
CR613169
LOCUS
DEFINITION
full-length cDNA clone CS0DE013YI20 of Placenta of Homo sapiens
(human) .
ACCESSION
CR613169
VERSION
CR613169.1 GI:50493976
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2716)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 2716)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
```



- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.

FEATURES  
Source  
1..2716  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE013YJ20"  
/tissue\_type="Placenta"  
/plasmid="pCMVSPORT\_6"

Query Match		9.6%; Score 731; DB 6; Length 2716;
Best Local Similarity		99.4%; Pred. No. 0;
Matches 1031; Conservative		0; Mismatches 6; Indels 0; Gaps 0;
QY	5685	AGAGTTTCTATGAGATGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744
Db	877	AGAGTTTCTATGAGATGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 936
QY	5745	TTSTAAGGGAAACCCACCTTCACTGCCACACCCATATGSCCGCAACTGCTATCACTC 5804
Db	937	TTCTAAGGNAACCCACCTTCACTGCCACACCCATATGSCCGCAACTGCTATCACTC 996
QY	5805	TGCCACTCTTTGCATGATGATCAATATCTATTGGACAGGAAATGATTAATCTAG 5864
Db	997	TGCCACTCTTTGCATGATGATCAATATCTATTGGACAGGAAATGATTAATCTAG 1056
QY	5865	TTGTCCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGC 5924
Db	1057	TTGTCCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGC 1116
QY	5925	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAAGAAATATCTCCCA 5984
Db	1117	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAAGAAATATCTCCCA 1176
QY	5985	ACTCACCGGGTACATGGGACCTCTAGCCCTTCAAGGACTAGTCTCTCAAACTACA 6044
Db	1177	ACTCACCGGGTACATGGGACCTCTAGCCCTTCAAGGACTAGTCTCTCAAACTACA 1236
QY	6045	TGAACCTCCGTACCCATATCGCTGTGTAAGCTTATTAATACCACTCACTGGGCT 6104
Db	1237	TGAACCTCCGTACCCATATCGCTGTGTAAGCTTATTAATACCACTCACTGGGCT 1296
QY	6105	CCATGAGTCTCGGCCCAACCCCTACTAACTGTGGATATGCTCCCTGAACTTCAR 6164
Db	1297	CCATGAGTCTCGGCCCAACCCCTACTAACTGTGGATATGCTCCCTGAACTTCAR 1356
QY	6165	GCCATATGTTTCAATCCCTGTACTGAAACAATGGAAACAACCTTCAGCACAGAAATAACAC 6224
Db	1357	GCCATATGTTTCAATCCCTGTACTGAAACAATGGAAACAACCTTCAGCACAGAAATAACAC 1416
QY	6225	CACCTCCGTTTATGAGGACCTCTGTTTCCAATSTGGAATAAACCCATACCTCAAACT 6284
Db	1417	CACCTCCGTTTATGAGGACCTCTGTTTCCAATCTGGAATAAACCCATACCTCAAACT 1476
QY	6285	CACCTGTGTAAATTTAGCAATACTACATACACCAACTCCCAATGATCAGGTGGGT 6344
Db	1477	CACCTGTGTAAATTTAGCAATACTACATACACCAACTCCCAATGATCAGGTGGGT 1536
QY	6345	AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6404
Db	1537	AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1596
QY	6405	AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATATGCTCTCTCTCATTTCTAGTGCC 6464
Db	1597	AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATATGCTCTCTCTCATTTCTAGTGCC 1656
QY	6465	CCCYATGRCATCTACATGAAACAAGATTTATACAGTTATGATATATCTTAAGCCCGCA 6524

Db	1657	CCCTATGACCATCTACACTGAACAAGATTATATAGTATGTATATCTAAGCCCGCAA 1716
QY	6525	CAAAAGAGTACCCATCTCTCTTTTGTATAGGAGAGGAGTGTAGGTGCACCTAGGTAC 6584
Db	1717	CAAAAGAGTACCCATCTCTCTTTTGTATAGGAGAGGAGTGTAGGTGCACCTAGGTAC 1776
QY	6585	TGGCATTGGCGGTATCACAACTCTACTAGTCTTACTACAACTATCTCAAGAACTAAA 6644
Db	1777	TGGCATTGGCGGTATCACAACTCTACTAGTCTTACTACAACTATCTCAAGAACTAAA 1836
QY	6645	TGGGGACATGGAACGGGTGCGGACCTCCCTGGTCACTTGCAGAGTCAACTTAATCCTCT 6704
Db	1837	TGGGGACATGGAACGGGTGCGGACCTCCCTGGTCACTTGCAGAGTCAACTTAATCCTCT 1896
QY	6705	AGCAGCAGTAGTCTCTTC 6721
Db	1897	AGCAGCAGTAGTCTCTTC 1913

CR605851 2748 bp mRNA linear HTC 21-JUL-2004  
full-length cDNA clone CS0DE012YJ24 of Placenta of Homo sapiens (human).  
ACCESSION CR605851  
VERSION CR605851.1 GI:50486658  
KEYWORDS HTC; CNSLT\_CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2748)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
REFERENCE 2 (bases 1 to 2748)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES  
Location/Qualifiers  
1..2748  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE012YJ24"  
/tissue\_type="Placenta"  
/plasmid="pCMVSPORT\_6"  
source

ORIGIN		Query Match 9.6%; Score 731; DB 6; Length 2748;
Best Local Similarity		99.4%; Pred. No. 0;
Matches 1031; Conservative		0; Mismatches 6; Indels 0; Gaps 0;
QY	5685	AGAGTTTCTATGAGATGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744
Db	876	AGAGTTTCTATGAGATGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 935
QY	5745	TTSTAAGGGAAACCCACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCCTC 5804
Db	936	TTCTAAGGGAAACCCACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCCTC 995



```
QY 6225 CACTTCGGTTTGTAGTAGACCTCTGTGTTTCAATSTGGAATACCCATACCTCCTCAACCT 6284
Db 1427 CACTTCGGTTTGTAGTAGACCTCTGTGTTTCAATSTGGAATACCCATACCTCCTCAACCT 1486
QY 6285 CACTTCGGTTTGTAGTAGACCTCTGTGTTTCAATSTGGAATACCCATACCTCCTCAGTGGGT 6344
Db 1487 CACTTCGGTTTGTAGTAGACCTCTGTGTTTCAATSTGGAATACCCATACCTCCTCAGTGGGT 1546
QY 6345 AACTTCCTCCACACAAATAGTCTGCTTACCTCAGGATATTTTGTCTGTGTGCTACCTC 6404
Db 1547 AACTTCCTCCACACAAATAGTCTGCTTACCTCAGGATATTTTGTCTGTGTGCTACCTC 1606
QY 6405 AGCTATCGTTTGTGTAATGGCTCTTCAGAAATCTATGTGCTCTCTCATTTCTAGTGCC 6464
Db 1607 AGCTATCGTTTGTGTAATGGCTCTTCAGAAATCTATGTGCTCTCTCATTTCTAGTGCC 1666
QY 6465 CCCATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524
Db 1667 CCCATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1726
QY 6525 CAAGAAGTAGTACCAATCTCTCTCTTTGTTATAGGAGGAGTGTAGGTGCATAGTAC 6584
Db 1727 CAAGAAGTAGTACCAATCTCTCTCTTTGTTATAGGAGGAGTGTAGGTGCATAGTAC 1786
QY 6585 TGGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1787 TGGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1846
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCT 6704
Db 1847 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCT 1906
QY 6705 AGCAGCAGTAGTCTTTC 6721
Db 1907 AGCAGCAGTAGTCTTTC 1923

RESULT 5
LOCUS CR617248
DEFINITION full-length cDNA clone CS0D1022YJ18 of Placenta HTC 21-JUL-2004
of Homo sapiens (human).
ACCESSION CR617248
VERSION CR617248.1 GI:50498055
KEYWORDS HTC; CNSLT CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2749)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2749)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..2749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

## ORIGIN

```
Query Match 9.6%; Score 731; DB 6; Length 2749;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGATTTCTATGAGGAATGAGCGTCCCGGAAATATTGATGCCCCCAATCGTATAGGAGTCT 5744
Db 881 AGATTTCTATGAGGAATGAGCGTCCCGGAAATATTGATGCCCCCAATCGTATAGGAGTCT 940
QY 5745 TTSTAAAGGNAACCCACACCTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 5804
Db 941 TTCTAAAGGNAACCCACACCTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 1000
QY 5805 TGCACCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5864
Db 1001 TGCACCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1060
QY 5865 TTGTCTCTGGAGGACTTTGGAGTCACTGTCTGTGTTGGAGTCTTCAACCCCAACTGCTATGTC 5924
Db 1061 TTGTCTCTGGAGGACTTTGGAGTCACTGTCTGTGTTGGAGTCTTCAACCCCAACTGCTATGTC 1120
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACAATGTAAGAAAGTAATCTCCCA 5984
Db 1121 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACAATGTAAGAAAGTAATCTCCCA 1180
QY 5985 ACTCACCSGGGTACATGGGACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACA 6044
Db 1181 ACTCACCSGGGTACATGGGACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACA 1240
QY 6045 TGAACCCCTCCGTACCCATACCTCGCTGGTAAAGCTATTAAATACCAACCTCTCTGGGCT 6104
Db 1241 TGAACCCCTCCGTACCCATACCTCGCTGGTAAAGCTATTAAATACCAACCTCTCTGGGCT 1300
QY 6105 CCATGAGGTCTGGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCCCTGAACCTTCAR 6164
Db 1301 CCATGAGGTCTGGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCCCTGAACCTTCAG 1360
QY 6165 GCCATATGTTCAATCCCTGTACCTGCAACCAATGGAACTTCCAGCACAGAAATAAACAC 6224
Db 1361 GCCATATGTTCAATCCCTGTACCTGCAACCAATGGAACTTCCAGCACAGAAATAAACAC 1420
QY 6225 CACTTCGGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCCTCAACCT 6284
Db 1421 CACTTCGGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCCTCAACCT 1480
QY 6285 CACTTCGTGTAATAATTTAGCAATATCTACATACACCAACCTCCCAATGCATCAGGTGGGT 6344
Db 1481 CACTTCGTGTAATAATTTAGCAATATCTACATACACCAACCTCCCAATGCATCAGGTGGGT 1540
QY 6345 AACTCTCCACACAAATAGTCTGCTACCTCAGGAAATATTTTGTCTGTGTGCTACCTC 6404
Db 1541 AACTCTCCACACAAATAGTCTGCTACCTCAGGAAATATTTTGTCTGTGTGCTACCTC 1600
QY 6405 AGCTATCGTTGTTGAAATGGCTCTTCAGAAATCTATGTGCTCTCTCATTTCTAGTGCC 6464
Db 1601 AGCTATCGTTGTTGAAATGGCTCTTCAGAAATCTATGTGCTCTCTCATTTCTAGTGCC 1660
QY 6465 CCCTATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524
Db 1661 CCCTATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1720
QY 6525 CAAAGAGTACCAATCTCTCTTTTGTGTTATAGGAGCAGGAGTGTAGGTGCATAGGTAC 6584
Db 1721 CAAAGAGTACCAATCTCTCTTTTGTGTTATAGGAGCAGGAGTGTAGGTGCATAGGTAC 1780
QY 6585 TGGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1781 TGGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1840
```

```
/clone="CS0D1022YJ18"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
```

```
QY 6645 TGGGACATGGAACGGGTGCGGACTCCTCGTGCACCTTGCAGATCAACTTAACCTCCT 6704
|||||
Db 1841 TGGGACATGGAACGGGTGCGGACTCCTCGTGCACCTTGCAGATCAACTTAACCTCCT 1900

QY 6705 AGCAGCAGTAGTCCTTC 6721
|||||
Db 1901 AGCAGCAGTAGTCCTTC 1917

RESULT 6
BX391741/c 935 bp mRNA linear EST 29-APR-2004
LOCUS BX391741 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS01051YM13 3-PRIME, mRNA sequence.
ACCESSION BX391741
VERSION BX391741.1 GI:30619529
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 935)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BA1022ZA10_CS02089_1&c=4215.r

FEATURES
source Location/Qualifiers
1..935
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01051YM13"
/tissue_lib="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 6.9%; Score 525; DB 4; Length 935;
Best Local Similarity 99.5%; Pred. No. 8.3e-282; Mismatches 4; Indels 0; Gaps 0;
Matches 725; Conservative 0;

QY 5993 GGGTACATGGCACCTTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAACCC 6052
|||||
Db 746 GGGTACATGGCACCTTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAACCC 687

QY 6053 TCCGTACCATACCTGCTGGTAAGCCTATTTAATACACCTCTACTGGGCTCCATGAGG 6112
|||||
Db 686 TCCGTACCATACCTGCTGGTAAGCCTATTTAATACACCTCTACTGGGCTCCATGAGG 627

QY 6113 TCTCGSCCAAAACCTACTAATGTTGATATGCTCCCTGCACTTCAGGCCATATG 6172
|||||
Db 626 TCTCGSCCAAAACCTACTAATGTTGATATGCTCCCTGCACTTCAGGCCATATG 567

QY 6173 TTTCAATCCCTGACTGTAACCAATGGAAACAACCTTCAGCACAGAAATAAACACCACTTCG 6232
|||||
Db 566 TTTCAATCCCTGACTGTAACCAATGGAAACAACCTTCAGCACAGAAATAAACACCACTTCG 507
```

```
6233 TTTTAGTAGGACCTCTTGTGTTTCCAAATGGAATAAACCCATCACTCAAACTCACTG 6292
|||||
Db 506 TTTTAGTAGGACCTCTTGTGTTTCCAAATGGAATAAACCCATCACTCAAACTCACTG 447
|||||
QY 6293 TAAATTTAGCAATATACATACACCAACTCCCAATGCATCAGGTGGGTAACCTC 6352
|||||
Db 446 TAAATTTAGCAATATACATACACCAACTCCCAATGCATCAGGTGGGTAACCTC 387
|||||
QY 6353 CCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATC 6412
|||||
Db 386 CCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATC 327
|||||
QY 6413 GTTGTTCGAATGGCTCTTCAGAACTATATGCTCTCTCACTTCTTAGTGCCCCVATGR 6472
|||||
Db 326 GTTGTTCGAATGGCTCTTCAGAACTATATGCTCTCTCACTTCTTAGTGCCCCVATGA 267
|||||
QY 6473 CCATCTACATGAACAAGATTATACAGTTATGTATCTTAAGCCCGCAACAAAAGAG 6532
|||||
Db 266 CCATCTACATGAACAAGATTATACAGTTATGTATCTTAAGCCCGCAACAAAAGAG 207
|||||
QY 6533 TACCATTTCTCTTTTGTATAGAGCAGAGTGCTAGGTGCATAGGTACTGCATTCG 6592
|||||
Db 206 TACCATTTCTCTTTTGTATAGAGCAGAGTGCTAGGTGCATAGGTACTGCATTCG 147
|||||
QY 6593 GCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTTAAATGGGACA 6652
|||||
Db 146 GCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTTAAATGGGACA 87
|||||
QY 6653 TGGAAAGCGGTGCGGACTCCCTGGTTCACCTTTCGAAGATCAACTTAACTCCCTAGCAGCAG 6712
|||||
Db 86 TGGAAAGCGGTGCGGACTCCCTGGTTCACCTTTCGAAGATCAACTTAACTCCCTAGCAGCAG 27
|||||
QY 6713 TAGTCCTTC 6721
|||||
Db 26 TAGTCCTTC 18

RESULT 7
BX367907 828 bp mRNA linear EST 28-APR-2004
LOCUS BX367907 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS01051YM13 5-PRIME, mRNA sequence.
ACCESSION BX367907
VERSION BX367907.1 GI:30459586
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 828)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AT022ZG01_T02023_1&c=4215.r.

FEATURES
source Location/Qualifiers
1..828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
```



QY 6285 CACCTGTGTAAATTTAGCAATACTACATACACCAACTCCCAATGCATCAGTGGGT 6344  
 |||||  
 Db 751 CACCTGTGTAAATTTAGCAATACTACATACACCAACTCCCAATGCATCAGTGGGT 810  
 |||||  
 QY 6345 AACTCTCCACACAAATAGTCTGCCTACCTCAGGA 6381  
 |||||  
 Db 811 AACTCTCCACACAAATAGTCTGCCTACCTCAGGA 847  
 |||||

RESULT 9  
 BX408733 883 bp mRNA linear EST 30-APR-2004  
 LOCUS BX408733 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012J24  
 DEFINITION 5-PRIME, mRNA sequence.

ACCESSION BX408733  
 VERSION BX408733.1 GI:30635957  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.

## REFERENCE

1 (bases 1 to 883)  
 Li W.B., Gruber C., Jessee J. and Polayes D.

Full-length cDNA libraries and normalization

## JOURNAL

Unpublished (2001)

## COMMENT

Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.

This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0BAA008ZE02\_CS00686\_1&c=4215.r

## FEATURES

source

Location/Qualifiers

1..883  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE012YJ24"  
 /tissue type="PLACENTA"  
 /clone lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

## ORIGIN

Query Match 5.8%; Score 437; DB 4; Length 883;  
 Best Local Similarity 99.4%; Pred. No. 2.2e-232;  
 Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGACGGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
 |||||  
 Db 75 AGAGTTTCTATGAGAAATGACGGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 134  
 |||||

QY 5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804  
 |||||  
 Db 135 TTCTAAGGGAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 194  
 |||||

QY 5805 TGGCACTCTTGCATGCATGCAAAATACTATTTCGACAGGAAAATGATTAATCCTAG 5864  
 |||||

Db 195 TGGCACTCTTGCATGCATGCAAAATACTATTTCGACAGGAAAATGATTAATCCTAG 254  
 |||||

QY 5865 TTGTCTCGGAGGACTTGGAGTCACTCTGTTCGACTTACTTCACCAAACTGGTATGTC 5924  
 |||||

Db 255 TTGTCTCGGAGGACTTGGAGTCACTCTGTTCGACTTACTTCACCAAACTGGTATGTC 314  
 |||||

QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAACTCTCCCA 5984  
 |||||  
 Db 315 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAACTCTCCCA 374  
 |||||

QY 5985 ACTCACCSGGGTACATGCGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 6044  
 |||||

Db 375 ACTCACCSGGGTACATGCGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 434  
 |||||

QY 6045 TGAACCCCTCGTACCCATCTCGCTCGTGAAGCCCTATTATATACCAACCCCTCACTGGGT 6104  
 |||||

Db 435 TGAACCCCTCGTACCCATCTCGCTCGTGAAGCCCTATTATATACCAACCCCTCACTGGGT 494  
 |||||

QY 6105 CCATGAGTCTCGGCCCAAAACCTTACTAACTGTGGATATGCCCTCCCGCTGAACCTTCA 6164  
 |||||

Db 495 CCATGAGTCTCGGCCCAAAACCTTACTAACTGTGGATATGCCCTCCCGCTGAACCTTCA 554  
 |||||

QY 6165 GCATATGTTTCAATCCCTGTACCTTGAACAATGGAACAACCTTCAGCAGAAATAAACAC 6224  
 |||||

Db 555 GCATATGTTTCAATCCCTGTACCTTGAACAATGGAACAACCTTCAGCAGAAATAAACAC 614  
 |||||

QY 6225 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCAATSTGGAATTAACCCATACCTCAAACT 6284  
 |||||

Db 615 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCAATSTGGAATTAACCCATACCTCAAACT 674  
 |||||

QY 6285 CACTGTGTAAATTTAGCAATACTACATACACCAACT 6325  
 |||||

Db 675 CACTGTGTAAATTTAGCAATACTACATACACCAACT 715  
 |||||

RESULT 10  
 BX326647/c 872 bp mRNA linear EST 07-APR-2004  
 LOCUS BX326647 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0DI051YM13 3-PRIME, mRNA sequence.

## ACCESSION

BX326647

## VERSION

BX326647.1 GI:30334578

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.

1 (bases 1 to 872)

Li W.B., Gruber C., Jessee J. and Polayes D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0BAD010ZC11\_AD00948\_1&c=4215.r

## FEATURES

source

Location/Qualifiers

1..872

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051YM13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was normalized."

## ORIGIN

Query Match 5.7%; Score 431; DB 4; Length 872;  
 Best Local Similarity 99.4%; Pred. No. 5.2e-229;  
 Matches 631; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6038 AACTACATGAACCCCTCCGTACCCATACCTCGCTGGTAAGCCTATTAAATACCACTCA 6097  
 DB 693 AACTACATGAACCCCTCCGTACCCATACCTCGCTGGTAAGCCTATTAAATACCACTCA 634

QY 6098 CTGGGCTCCATAGAGTCTCGGCCCAAAACCCCTACTAACTGTGTGATGCTCCCTCGA 6157  
 DB 633 CTGGGCTCCATAGAGTCTCGGCCCAAAACCCCTACTAACTGTGTGATGCTCCCTCGA 574

QY 6158 ACTTCARGCCATATGTTTCAATCCCTGCTGAACATGGAACCACTTCAGCACAGAAA 6217  
 DB 573 ACTTCARGCCATATGTTTCAATCCCTGCTGAACATGGAACCACTTCAGCACAGAAA 514

QY 6218 TAAACACCACTTCGCTTTTGTAGTAGGACCTCTGTTCCTCAATSTGGAATTAACCCATCCT 6277  
 DB 513 TAAACACCACTTCGCTTTTGTAGTAGGACCTCTGTTCCTCAATCTGGAATTAACCCATCCT 454

QY 6278 CAAACCTCAGCTGTGTAATAATTTAGCAATACATACACAACTCCCAATGCATCA 6337  
 DB 453 CAAACCTCAGCTGTGTAATAATTTAGCAATACATACACAACTCCCAATGCATCA 394

QY 6338 GGTGGGTAACTCTCCACAAATAGTCTGCTACCTCAGCAATATTTTGTCTGTG 6397  
 DB 393 GGTGGGTAACTCTCCACAAATAGTCTGCTACCTCAGCAATATTTTGTCTGTG 334

QY 6398 GTACCTCAGCTATCGTTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATCT 6457  
 DB 333 GTACCTCAGCTATCGTTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATCT 274

QY 6458 TAGTGGCCCTATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCAATCTAAGC 6517  
 DB 273 TAGTGGCCCTATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCAATCTAAGC 214

QY 6518 CCGCAACAAAGAGTACCACTTCCTCTTTGTTATAGAGCAGGAGTCTAGGTGCAC 6577  
 DB 213 CCGCAACAAAGAGTACCACTTCCTCTTTGTTATAGAGCAGGAGTCTAGGTGCAC 154

QY 6578 TAGGTACTGCTATGCGGTATCACAACTCTACTCAGTTCTACTCAAACTATCTCAAG 6637  
 DB 153 TAGGTACTGCTATGCGGTATCACAACTCTACTCAGTTCTACTCAAACTATCTCAAG 94

QY 6638 AACTAAATGGGACATGGAACGGGTGCGGACTCC 6672  
 DB 93 AACTAAATGGGACATGGAACGGGTGCGGACTCC 59

RESULT 11  
 BX389656  
 LOCUS  
 DEFINITION  
 BX389656 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1051YML3 5-PRIME, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BX389656 931 bp mRNA linear EST 29-APR-2004  
 BX389656 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1051YML3 5-PRIME, mRNA sequence.

EST.  
 BX389656.1 GI:30462930  
 Homo sapiens (human)  
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 931)  
 Li, W. B., Gruber, C., Jessee, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 4215.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?cs=CS0BAF0272B05\_AF02531\_1&c=4215.r

FEATURES  
 source  
 1..931  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1051YML3"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 5.4%; Score 411; DB 4; Length 931;  
 Best Local Similarity 99.3%; Pred. No. 9e-218;  
 Matches 611; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5853 GATTAACTCCTAGTGTGCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCA 5912  
 DB 7 GATTAACTCCTAGTGTGCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCA 66

QY 5913 AACTGTATGTCTGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAACATGTAAGA 5972  
 DB 67 AACTGTATGTCTGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAACATGTAAGA 126

QY 5973 AGTAATCTCCCACTCACCSCGGTACATGCACCTCTAGCCCTACAAAGAGACTAGATCT 6032  
 DB 127 AGTAATCTCCCACTCACCSCGGTACATGCACCTCTAGCCCTACAAAGAGACTAGATCT 186

QY 6033 CTCAAACTACATGAACCCCTCCGTACCCATCTCGCTGGTAAGCCTATTTAATACCAC 6092  
 DB 187 CTCAAACTACATGAACCCCTCCGTACCCATCTCGCTGGTAAGCCTATTTAATACCAC 246

QY 6093 CCTCACTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTGGATATGCTTCC 6152  
 DB 247 CCTCACTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTGGATATGCTTCC 306

QY 6153 CTGGAATCTCARGCCATATGTTTCAATCCCTGACCTGACCAATGGAACCACTTCAGCAC 6212  
 DB 307 CTGGAATCTCARGCCATATGTTTCAATCCCTGACCTGACCAATGGAACCACTTCAGCAC 366

QY 6213 AGAAATAAACACACCTTCCTGTTTGTAGTAGGACCTCTGTTTCCCAATSTGGAATAACCCA 6272  
 DB 367 AGAAATAAACACACCTTCCTGTTTGTAGTAGGACCTCTGTTTCCCAATSTGGAATAACCCA 426

QY 6273 TACCTCAAAACCTCAGCTGTGTAATAATTTAGCAATACTACATACAAACCACTCCCAATG 6332  
 DB 427 TACCTCAAAACCTCAGCTGTGTAATAATTTAGCAATACTACATACAAACCACTCCCAATG 486

QY 6333 CATCAGGTGGGTAACTCTCCCAACAAATAGTCTGCTTACCTCAGGAATATTTTGT 6392  
 DB 487 CATCAGGTGGGTAACTCTCCCAACAAATAGTCTGCTTACCTCAGGAATATTTTGT 546

QY 6393 CTGTGTACTCTCAGCTATCTGTTTGAATGGCTCTTCAGAAATCTATGTCTCTCTC 6452  
 DB 547 CTGTGTACTCTCAGCTATCTGTTTGAATGGCTCTTCAGAAATCTATGTCTCTCTC 606

QY 6453 ATTCTTAGTGCCCCC 6467  
 DB 607 ATTCTTAGTGCCCCC 621

RESULT 12  
 DB380020  
 LOCUS

792 bp mRNA linear EST 13-NOV-2005





```
QY 5872 GGAGGACTTGGAGTCTGCTGTGGACTTACTTACCCAAACTGGTATGTCGTATGGG 5931
Db 61 GGAGGACTTGGAGTCTGCTGTGGACTTACTTACCCAAACTGGTATGTCGTATGGG 120
QY 5932 GTGGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAAAGTAATCTCCCACTCACC 5991
Db 121 GTGGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAAAGTAATCTCCCACTCACC 180
QY 5992 SGGGTACATGGCACTTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAGAAC 6051
Db 181 CGGGTACATGGCACTTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAGAAC 240
QY 6052 CTCGGTACCCATCTCGCTCGTGAAGCTATTTAATACCAACCTCTCACTGGGCTCCATGAG 6111
Db 241 CTCGGTACCCATCTCGCTCGTGAAGCTATTTAATACCAACCTCTCACTGGGCTCCATGAG 300
QY 6112 GTCTCGGCCAAAACCTTACTAATCTGTGGATATGCTCCCTCGAATCTTCARGCCATAT 6171
Db 301 GTCTCGGCCAAAACCTTACTAATCTGTGGATATGCTCCCTCGAATCTTCARGCCATAT 360
QY 6172 GTTTCATCCCTGTACTGAACATGAAGCAACTTCAGCACAGAAATAAACCACTTCC 6231
Db 361 GTTTCATCCCTGTACTGAACATGAAGCAACTTCAGCACAGAAATAAACCACTTCC 420
QY 6232 GTTTTAGTAGGACCTCTGTTTCCAAATSTGGAATAACCACTTCAACCTCACTCT 6291
Db 421 GTTTTAGTAGGACCTCTGTTTCCAAATCTGGAATAACCACTTCAACCTCACTCT 480
QY 6292 GTAAATTTAGCAATACATACACACAACTCCCAATGATCAGGTGGG 6343
Db 481 GTAAATTTAGCAATACATACACACAACTCCCAATGATCAGGTGGG 532
```

```
RESULT 14
BX347314
LOCUS
DEFINITION
BX347314 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CS0D1044YK06 5-PRIME, mRNA sequence.
BX347314
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
BX347314.1 GI:30365192
Homo sapiens (human)
Homo sapiens
ORGANISM
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with NotI and cloned  
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4215.r

For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?c=CS0BAA016ZH06\\_CS01489\\_1&c=4215.r](http://www.genoscope.cns.fr/cdna?c=CS0BAA016ZH06_CS01489_1&c=4215.r)

## FEATURES

## source

1. 924  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1044YK06"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was

## ORIGIN

Query Match 5.0%; Score 377; DB 4; Length 924;  
Best Local Similarity 99.6%; Pred. No. 1.1e-198;  
Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGCAGCGTCCCGGAAATATTGATGCCCAATCGTATAGGAGTCT 5744  
Db 216 AGAGTTTCTATGAGAAATGCAGCGTCCCGGAAATATTGATGCCCAATCGTATAGGAGTCT 275  
QY 5745 TTSTAAGGAAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 5804  
Db 276 TTSTAAGGAAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 335  
QY 5805 TGCCACTCTTTCATGATGCAATTAATCTATTTATGGACAGAAATATTAATCCTAG 5864  
Db 336 TGCCACTCTTTCATGATGCAATTAATCTATTTATGGACAGAAATATTAATCCTAG 395  
QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC 5924  
Db 396 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC 455  
QY 5925 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAGAAAACATGTAAAGAAAGTAATCTCCCA 5984  
Db 456 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAGAAAACATGTAAAGAAAGTAATCTCCCA 515  
QY 5985 ACTCACCCGGGTACATGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 6044  
Db 516 ACTCACCCGGGTACATGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 575  
QY 6045 TGAAACCCCTCCGTACCCATCTCGCTCGTGAAGCCCTATTATATCCACCCTCACTGGGT 6104  
Db 576 TGAAACCCCTCCGTACCCATCTCGCTCGTGAAGCCCTATTATATCCACCCTCACTGGGT 635  
QY 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCGAATTC 6163  
Db 636 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCGAATTC 694

## RESULT 15

DA838045  
LOCUS  
DEFINITION  
DA838045 PLACE6 Homo sapiens cDNA clone PLACE6004226 5', mRNA  
sequence.  
DA838045  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
EST.  
DA838045.1 GI:81276957  
Homo sapiens (human)  
Homo sapiens  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,  
Yamaoka, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakeguri, H.,  
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,  
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,  
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,  
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.  
Identification of Transcriptional Modulation: Large-scale  
Identification and Characterization of Putative Alternative  
Promoters of Human Genes  
Genome Res. 16 (1), 55-65 (2006)  
16344560  
Contact: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com

## TITLE

## JOURNAL

## PUBMED

## COMMENT

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES  
source  
1..538  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE6004226"  
/tissue\_type="placenta"  
/clone\_lib="PLACE6"  
/note="Vector: pME18SFL3"

ORIGIN

Query Match 4.9%; Score 374; DB 9; Length 538;  
Best Local Similarity 99.6%; Pred. No. 5,3e-197; Mismatches 2; Indels 0; Gaps 0;  
Matches 474; Conservative 0;  
QY 5698 GTTCTATGGAGAATGCAGCGTCGCGAAATATTCATGCCCATCGTATAGGAGTCTTTS 5747  
Db 1 GTTCTATGGAGAATGCAGCGTCGCGAAATATTCATGCCCATCGTATAGGAGTCTTTC 60  
QY 5748 TAAAGGAACCCACCTTCACCTGCGCCACACCCATATGCCCGCACTGCTATCACTCTGC 5807  
Db 61 TAAAGGAACCCACCTTCACCTGCGCCACACCCATATGCCCGCACTGCTATCACTCTGC 120  
QY 5808 CACTCTTTGCATGCATGCAATACATCTATTATGGACAGAAATGATTAATCCTAGTTG 5867  
Db 121 CACTCTTTGCATGCATGCAATACATCTATTATGGACAGAAATGATTAATCCTAGTTG 180  
QY 5868 TCCTGGAGGACCTGGAGTCACCTGTCTGTGGACTTACTTCCCAAACTGGTATGTCTGA 5927  
Db 181 TCCTGGAGGACCTGGAGTCACCTGTCTGTGGACTTACTTCCCAAACTGGTATGTCTGA 240  
QY 5928 TGGGGGTGGAGTCAAGATCAGGCAAGAGAAACATGTAAAGAGTAATCTCCCACT 5987  
Db 241 TGGGGGTGGAGTCAAGATCAGGCAAGAGAAACATGTAAAGAGTAATCTCCCACT 300  
QY 5988 CACCGGGTACATGGCAGCTCTAGCCCTACAAAGAGCTAGATCTCTCAAACTACATGA 6047  
Db 301 CACCGGGTACATGGCAGCTCTAGCCCTACAAAGAGCTAGATCTCTCAAACTACATGA 360  
QY 6048 AACCTCCGTACCCATCTCGCTGGTAAGCCCTATTTAATACCACCCCTCACTGGGCTCCA 6107  
Db 361 AACCTCCGTACCCATCTCGCTGGTAAGCCCTATTTAATACCACCCCTCACTGGGCTCCA 420  
QY 6108 TGAGGTCTCGGCCCAAAACCTACTAACTGTGGATATGCTCCCTCGAACTTCA 6163  
Db 421 TGAGGTCTCGGCCCAAAACCTACTAACTGTGGATATGCTCCCTCGAACTTCA 476

RESULT 16  
AUI138405  
LOCUS AUI138405 PLACE1 Homo sapiens cDNA clone PLACE1008489 5', mRNA linear EST 16-NOV-2004  
DEFINITION sequence.  
ACCESSION AUI138405  
VERSION AUI138405.2 GI:55780428  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo  
1 (bases 1 to 721)  
Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,  
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,  
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,  
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,  
Murakawa,K., Ishida,S., Ishibaashi,T., Takahashi-Fujii,A.,

Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.  
Diversification of Transcriptional Modulation: Large-scale  
Identification and Characterization of Putative Alternative  
Promoters of Human Genes  
Genome Res. 16 (1), 55-65 (2006)  
16344560  
On Oct 25, 2000 this sequence version replaced gi:10999926.  
Contact: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES  
source

1..721  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE1008489"  
/tissue\_type="placenta"  
/clone\_lib="PLACE1"  
/note="Vector: pME18SFL3"

ORIGIN

Query Match 4.9%; Score 373; DB 1; Length 721;  
Best Local Similarity 99.4%; Pred. No. 2e-196; Mismatches 0; Indels 0; Gaps 0;  
Matches 523; Conservative 0;  
QY 6165 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACCTTCAGCAGAGAAATAAACAC 6224  
Db 23 GCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACCTTCAGCAGAGAAATAAACAC 82  
QY 6225 CACTTCGGTTTATAGGACCTCTGTTCCAATSTGGAAATACCCATCTCAAACT 6284  
Db 83 CACTTCGGTTTATAGGACCTCTGTTCCAATSTGGAAATACCCATCTCAAACT 142  
QY 6285 CACTGTGTAAATTTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGT 6344  
Db 143 CACTGTGTAAATTTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGT 202  
QY 6345 AACTCCTCCACACAATAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404  
Db 203 AACTCCTCCACACAATAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 262  
QY 6405 AGCCTATCGTTTGTGAATGGCTCTTCAGATCTATGTGCTTCCTCTCATTTCTTAGTGCC 6464  
Db 263 AGCCTATCGTTTGTGAATGGCTCTTCAGATCTATGTGCTTCCTCTCATTTCTTAGTGCC 322  
QY 6465 CCCATGRCCTATCTACACTGAACAAGATTTATACAGTTATGTATATCTAAGCCCCGCAA 6524  
Db 323 CCCTATGACCACTACACTGAACAAGATTTATACAGTTATGTATATCTAAGCCCCGCAA 382  
QY 6525 CAAAGAGTACCCATCTTCTTTTGTATAGGACGAGGTGTGTAGTGCACTAGGTAC 6584  
Db 383 CAAAGAGTACCCATCTTCTTTTGTATAGGACGAGGTGTGTAGTGCACTAGGTAC 442  
QY 6585 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAGAACTAAA 6644  
Db 443 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAGAACTAAA 502  
QY 6645 TGGGACATGAACGGGTGCGGACTCCCTGGTCACTTGTCAAGAT 6690  
Db 503 TGGGACATGAACGGGTGCGGACTCCCTGGTCACTTGTCAAGAT 548

RESULT 17  
DA832031  
LOCUS DA832031 761 bp mRNA linear EST 13-NOV-2005

```

DEFINITION   DA832031 PLACE1 Homo sapiens cDNA clone PLACE1008285 5', mRNA
sequence.
ACCESSION    DA832031
VERSION      DA832031.1 GI:82376942
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE    1. (bases 1 to 761)
AUTHORS      Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE        Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL      Genome Res. 16 (1), 55-65 (2006)
PUBMED       16344560
COMMENT      Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute (HRI); cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
HRI.

FEATURES             Location/Qualifiers
     source           1..761
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1008285"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

ORIGIN
Query Match      4.9%; Score 373; DB 9; Length 761;
Best Local Similarity 99.2%; Pred. No. 2e-196;
Matches 623; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5835 TTATTGGACAGGAAAATGATTAACTCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTG 5894
DB 1 TTATTGGACAGGAAAATGATTAACTCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTG 60

QY 5895 TTGACTTACTTCAACCAACTGGTATGTCGTATGGGGTGGAGTTCAAGTCAAGCAAG 5954
DB 61 TTGACTTACTTCAACCAACTGGTATGTCGTATGGGGTGGAGTTCAAGTCAAGCAAG 120

QY 5955 AGAAAAATGTTAAAGAAAGTAACTCTCCAACTCACCAGGTTACATGGCACTCTAGCCCC 6014
DB 121 AGAAAAATGTTAAAGAAAGTAACTCTCCAACTCACCAGGTTACATGGCACTCTAGCCCC 180

QY 6015 CTCAAAAGGACTAGATCTCTCAAAAATACATGAAGAACCCCTCGTACCATCTCGCTGGT 6074
DB 181 CTCAAAAGGACTAGATCTCTCAAAAATACATGAAGAACCCCTCGTACCATCTCGCTGGT 240

QY 6075 AAGCCTATTTAATACCAACCTCACTGGTCCATGAGTCTCGGCCCAAAACCTACTAA 6134
DB 241 AAGCCTATTTAATACCAACCTCACTGGTCCATGAGTCTCGGCCCAAAACCTACTAA 300

QY 6135 CTGTTGGATGCTCCCTCGAAGTCTCARGCCATATGTTCAATCCCTGTACTCTGAACA 6194
DB 301 CTGTTGGATGCTCCCTCGAAGTCTCARGCCATATGTTCAATCCCTGTACTCTGAACA 360

QY 6195 ATGGAACAACCTCAGCACAGAAAATAACCACTTCCGTTTATAGTAGGACCTCTGTGTTTC 6254

```

```

Db 361 ATGGAACAACCTCAGCACAGAAAATAGACACCACCTCCGTTTATAGTAGGACCTCTGTGTTTC 420
QY 6255 CAATSTGGAATAAACCCATACCTCAAAACCTCAGCTGTGTAAATAATTTAGCAATACTACATA 6314
Db 421 CAATCTGGAATAAACCCATACCTCAAAACCTCAGCTGTGTAAATAATTTAGCAATACTACATA 480
QY 6315 CACAACCAACTCCCAATGCATCAGGTGGGTAATCTCTCCACACAAATAGTGTGCTTACC 6374
Db 481 CACAACCAACTCCCAATGCATCAGGTGGGTAATCTCTCTCCACACAAATAGTGTGCTTACC 540
QY 6375 CTCAGAATAATTTTGTCTGTGTGACCTCAGCTATCGTTGTTGATGGCTCTTACAGA 6434
Db 541 CTCAGAATAATTTTGTCTGTGTGACCTCAGCTATCGTTGTTGATGGCTCTTACAGA 600
QY 6435 ATCTATGTGCTCTCTCTCATTTCTTAGTG 6462
Db 601 ATCTATGTGCTCTCTCTCATTTCTTAGTG 628

RESULT 18
DA849357
LOCUS       DA849357 PLACE6 Homo sapiens cDNA clone PLACE6018680 5', mRNA
DEFINITION sequence.
ACCESSION   DA849357
VERSION     DA849357.1 GI:82140231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE   1. (bases 1 to 540)
AUTHORS     Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE        Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL      Genome Res. 16 (1), 55-65 (2006)
PUBMED       16344560
COMMENT      Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB

FEATURES             Location/Qualifiers
     source           1..540
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE6018680"
                     /tissue_type="placenta"
                     /clone_lib="PLACE6"
                     /note="Vector: pME18SFL3"

ORIGIN
Query Match      4.9%; Score 372; DB 9; Length 540;
Best Local Similarity 99.6%; Pred. No. 7e-196;
Matches 472; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 5685 AGAGTTTCTATGAGAAATGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
67 AGAGTTTCTATGAGAAATGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 126  
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCCATATGCCCCGCAATGCTATCACTC 5804  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
127 TTCTAAGGGAACCCCACTTCACTGCCCCACACCCCATATGCCCCGCAATGCTATCACTC 186  
QY 5805 TGCCACTCTTTGATGATGCAAAATCACTATTATTGGACAGAGAAAATGATTAATCCCTAG 5864  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
187 TGCCACTCTTTGATGATGCAAAATCACTATTATTGGACAGAGAAAATGATTAATCCCTAG 246  
QY 5865 TTGTCCTGAGGACTTTGGAGTCACTGCTGTTGGACTTACTTCAACCACCACTGGTATGTC 5924  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
247 TTGTCCTGAGGACTTTGGAGTCACTGCTGTTGGACTTACTTCAACCACCACTGGTATGTC 306  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAAGTAATCTCCCA 5984  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
307 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAAGTAATCTCCCA 366  
QY 5985 ACTACCCGGGTACATGGCACTTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 6044  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
367 ACTACCCGGGTACATGGCACTTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 426  
QY 6045 TGAACCCCTCCGTACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6104  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
427 TGAACCCCTCCGTACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486  
QY 6105 CCATGAGGTCTCGGGCCCAAAACCTACTAACTGTTGGATATGCTTCCCTCCCTGAA 6158  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
487 CCATGAGGTCTCGGGCCCAAAACCTACTAACTGTTGGATATGCTTCCCTCCCTGAA 540

RESULT 19  
DA830049  
LOCUS DA830049 PLACE1 Homo sapiens cDNA clone PLACE1004321 5', mRNA  
DEFINITION DA830049  
ACCESSION DA830049.1 GI:82330584  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 819)  
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,  
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,  
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, K., Kushida, N.,  
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,  
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,  
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.  
Diversification of Transcriptional Modulation: Large-scale  
Identification and Characterization of Putative Alternative  
Promoters of Human Genes  
Genome Res. 16 (1), 55-65 (2006)  
16344560  
Contact: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com  
HRI human cDNA project: 5' - & 3'-end one pass sequencing: Helix  
Research Institute (HRI); cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
HRI.

TITLE  
JOURNAL  
PUBMED  
COMMENT

CONTACT: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com  
HRI human cDNA project: 5' - & 3'-end one pass sequencing: Helix  
Research Institute (HRI); cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
HRI.  
Location/Qualifiers  
1. .819  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="PLACE1004321"  
/issue\_type="placenta"  
/clone\_lib="PLACE1"  
/note="Vector: pME18SFL3"  
ORIGIN  
Query Match 4.9%; Score 370; DB 9; Length 819;  
Best Local Similarity 99.3%; Pred. No. 9.6e-195;  
Matches 570; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 5685 AGAGTTTCTATGAGAAATGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
Db 68 AGAGTTTCTATGAGAAATGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 127  
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCCATATGCCCCGCAATGCTATCACTC 5804  
Db 128 TTCTAAGGGAACCCCACTTCACTGCCCCACACCCCATATGCCCCGCAATGCTATCACTC 187  
QY 5805 TGCCACTCTTTGATGATGCAAAATCACTATTATTGGACAGAGAAAATGATTAATCCCTAG 5864  
Db 188 TGCCACTCTTTGATGATGCAAAATCACTATTATTGGACAGAGAAAATGATTAATCCCTAG 247  
QY 5865 TTGTCCTGAGGACTTTGGAGTCACTGCTGTTGGACTTACTTCAACCACCACTGGTATGTC 5924  
Db 248 TTGTCCTGAGGACTTTGGAGTCACTGCTGTTGGACTTACTTCAACCACCACTGGTATGTC 307  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAAGTAATCTCCCA 5984  
Db 308 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAAGTAATCTCCCA 367  
QY 5985 ACTACCCGGGTACATGGCACTTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 6044  
Db 368 ACTACCCGGGTACATGGCACTTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 427  
QY 6045 TGAACCCCTCCGTACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6104  
Db 428 TGAACCCCTCCGTACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487  
QY 6105 CCATGAGGTCTCGGGCCCAAAACCTACTAACTGTTGGATATGCTTCCCTCCCTGAA 6164  
Db 488 CCATGAGGTCTCGGGCCCAAAACCTACTAACTGTTGGATATGCTTCCCTCCCTGAA 547  
QY 6165 GCCATATGTTTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6224  
Db 548 GCCATATGTTTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607  
QY 6225 CACTTCCCTTTTAGTAGGACCTCTTGTTCCTCAAT 6258  
Db 608 CACTTCCCTTTTAGTAGGACCTCTTGTTCCTCAAT 641

RESULT 20  
DA8300734  
LOCUS DA8300734  
DEFINITION DA8300734 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012124  
5-PRIME, mRNA sequence.  
ACCESSION DA8300734  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 921)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 4215.1

For more information about this cluster, see [http://www.genoscope.cns.fr/cdna?cs=CS0BAA0082E02\\_CS00686\\_2&c=4215.1](http://www.genoscope.cns.fr/cdna?cs=CS0BAA0082E02_CS00686_2&c=4215.1)

#### FEATURES

Location/Qualifiers

1..921

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DF012XJ24"

/tissue\_type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA"

/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

#### ORIGIN

Query Match 4.9%; Score 369; DB 4; Length 921;

Best Local Similarity 99.6%; Pred. No. 3.5e-194;

Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5748 TAAGGGAACCCCACTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGC 5807

Db |||||

139 TAAGGGAACCCCACTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGC 198

QY 5808 CACTCTTTGATGATGCAAAATCTATTTATTTGGACAGGAAAAATGATTAATCCTAGTTG 5867

Db |||||

199 CACTCTTTGATGATGCAAAATCTATTTATTTGGACAGGAAAAATGATTAATCCTAGTTG 258

QY 5868 TCCTGGAGGACTTGGAGTCACTCTGTGGACTTACTTCAACCAACTGGTATGCTGA 5927

Db |||||

259 TCCTGGAGGACTTGGAGTCACTCTGTGGACTTACTTCAACCAACTGGTATGCTGA 318

QY 5928 TGGGGGTGGAGTTCAAGATCAGCAAGAGAAAAACATGTAAGAAAGTAACTCCCAACT 5987

Db |||||

319 TGGGGGTGGAGTTCAAGATCAGCAAGAGAAAAACATGTAAGAAAGTAACTCCCAACT 378

QY 5988 CACCSGGGTATGCGACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGA 6047

Db |||||

379 CACCSGGGTATGCGACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGA 438

QY 6048 AACCTTCGCTACCCACTTACGCTTGAAGCTTATTAATACACCTTCACTGGGCTCCA 6107

Db |||||

439 AACCTTCGCTACCCACTTACGCTTGAAGCTTATTAATACACCTTCACTGGGCTCCA 498

QY 6108 TGAGGTCTCGGCCCAAAACCTACTAACTTGTGGATATGCTCCCTCGAATTCARGCC 6167

Db |||||

499 TGAGGTCTCGGCCCAAAACCTACTAACTTGTGGATATGCTCCCTCGAATTCARGCC 558

QY 6168 ATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAATTCAGCACAGAAAT 6218

Db |||||

559 ATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAATTCAGCACAGAAAT 609

QY |||||

RESULT 21

DA846799

LOCUS

DEFINITION

DA846799 PLACE6 Homo sapiens cDNA clone PLACE6015377 5', mRNA

sequence.

DA846799

VERSION

DA846799.1 GI:82041683

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

#### REFERENCE

##### AUTHORS

Hominidae; Homo.  
1 (bases 1 to 571)  
Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,  
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,  
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,  
Yoneyama,T., Otsuka,B., Kanda,K., Yokoi,T., Kondo,H., Wagatsuna,M.,  
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,  
Tanabe,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.  
Diversification and Characterization of Putative Alternative  
Promoters of Human Genes  
Genome Res. 16 (1), 55-65 (2006)

#### JOURNAL

##### PUBMED

##### COMMENT

Contact: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology  
Developmental Organization, Japan); cDNA library construction:  
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,  
Research Association for Biotechnology (RAB) and Biotechnology  
Center, National Institute of Technology and Evaluation; 3'-end one  
pass sequencing: RAB.

#### FEATURES

##### source

##### Location/Qualifiers

##### 1..571

##### /organism="Homo sapiens"

##### /mol\_type="mRNA"

##### /db\_xref="taxon:9606"

##### /clone="PLACE6015377"

##### /tissue\_type="placenta"

##### /clone\_lib="PLACE6"

##### /note="Vector: pME18SFL3"

##### ORIGIN

##### Query Match

##### Best Local Similarity

##### Matches

##### Conservative

##### 0; Mismatches

##### 4; Indels

##### 0; Gaps

##### 0;

##### QY

##### 6003

##### CACTCTAGCCCCCTCAAGGACTAGATCTCTCAAAACTACATGAACCTCGTACCCA

##### 6062

##### Db

##### 1

##### CACTCTAGCCCCCTCAAGGACTAGATCTCTCAAAACTACATGAACCTCGTACCCA

##### 60

##### QY

##### 6063

##### TACTGCGCTGTGAAGCTATTAAATACACACCTCTCACTGGGCTCCATGAGGTCTCGGCCCA

##### 6122

##### Db

##### 61

##### TACTGCGCTGTGAAGCTATTAAATACACACCTCTCACTGGGCTCCATGAGGTCTCGGCCCA

##### 120

##### QY

##### 6123

##### AAACCTTACTAACTGTGGATATGCTCCCTCGAATTCARGCCATATGTTTCAATCCC

##### 6182

##### Db

##### 121

##### AAACCTTACTAACTGTGGATATGCTCCCTCGAATTCARGCCATATGTTTCAATCCC

##### 180

##### QY

##### 6183

##### TGTACTGAACTGAACCACTTCTAGGACAGAAATAAACACACTTCGCTTTTAGTAGG

##### 6242

##### Db

##### 181

##### TGTACTGAACTGAACCACTTCTAGGACAGAAATAAACACACTTCGCTTTTAGTAGG

##### 240

##### QY

##### 6243

##### ACCTCTGTTTCCCAATSTGGAATAACCCATACCTCAACACCTCACTGTGTAATTTAG

##### 6302

##### Db

##### 241

##### ACCTCTGTTTCCCAATSTGGAATAACCCATACCTCAACACCTCACTGTGTAATTTAG

##### 300

##### QY

##### 6303

##### CAATCTACATACACCAACTCCCAATGATCAGGTGGGTAACTCTCCACACCAAT

##### 6362

##### Db

##### 301

##### CAATCTACATACACCAACTCCCAATGATCAGGTGGGTAACTCTCCACACCAAT

##### 360

##### QY

##### 6363

##### AGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTGAA

##### 6422

##### Db

##### 361

##### AGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTGAA

##### 420

##### QY

##### 6423

##### TGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGGCCCTTATGACCATCTACAC

##### 6482

##### Db

##### 421

##### TGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGGCCCTTATGACCATCTACAC

##### 480

QY 6483 TGACACAAGATTTATACAGTTATGTCTATATCTAAGCCCGCAACAAAAGAGTACCCATTCT 6542  
 |||||  
 Db 481 TGACACAAGATTTATACAGTTATGTCTATATCTAAGCCCGCAACAAAAGAGTACCCATTCT 540  
 |||||  
 QY 6543 TCCTTTTGTATAGGAGCAGGAGTGCTAGGT 6573  
 |||||  
 Db 541 TCCTTTTGTATAGGAGCAGGAGTGCTAGGT 571  
 |||||

RESULT 22  
 DA835080  
 LOCUS DA835080 PLAC66 Homo sapiens cDNA clone PLACE6000237 5', mRNA  
 DEFINITION sequence.  
 ACCESSION DA835080  
 VERSION DA835080.1 GI:83059534  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 541)  
 AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,  
 Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,  
 Ishii, S., Sugiyama, T., Saio, K., Isono, Y., Irie, R., Kushida, N.,  
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,  
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,  
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.  
 TITLE Diversification of Transcriptional Modulation: Large-scale  
 Identification and Characterization of Putative Alternative  
 Promoters of Human Genes  
 JOURNAL Genome Res. 16 (1), 55-65 (2006)  
 PUBMED 16344560  
 COMMENT Contact: Takao Isogai  
 FLJ Project (HRI Team)  
 Helix Research Institute  
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology  
 Developmental Organization, Japan); cDNA library construction:  
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,  
 Research Association for Biotechnology (RAB) and Biotechnology  
 Center, National Institute of Technology and Evaluation; 3'-end one  
 pass sequencing: RAB.  
 FEATURES Location/Qualifiers  
 source 1..541  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="PLACE6000237"  
 /tissue\_type="placenta"  
 /clone\_lib="PLACE6"  
 /note="Vector: pME18SFL3"

ORIGIN  
 Query Match 4.8%; Score 363; DB 9; Length 541;  
 Best Local Similarity 99.6%; Pred. No. 8e-191;  
 Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 5685 AGAGTTTCTATGAGAAATCGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
 |||||  
 Db 77 AGAGTTTCTATGAGAAATCGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 136  
 |||||

QY 5745 TTSTAAGGAACCCCACTTCACTGCCCAACCCCATATGCCCGCACTGCTATCACTC 5804  
 |||||  
 Db 137 TTSTAAGGAACCCCACTTCACTGCCCAACCCCATATGCCCGCACTGCTATCACTC 196  
 |||||  
 QY 5805 TGCCACTCTTTGCATGCAAAATCTCATTATTGGACAGGAAAAATGATTAATCCCTAG 5864  
 |||||  
 Db 197 TGCCACTCTTTGCATGCAAAATCTCATTATTGGACAGGAAAAATGATTAATCCCTAG 256  
 |||||

QY 5865 TTGTCCCTGGAGGACTTGGAGTCACTGTCTGTGTGGACTTACTTCACCCAAACTGGTATGTC 5924  
 |||||  
 Db 257 TTGTCCCTGGAGGACTTGGAGTCACTGTCTGTGTGGACTTACTTCACCCAAACTGGTATGTC 316  
 |||||  
 QY 5925 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCA 5984  
 |||||  
 Db 317 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCA 376  
 |||||  
 QY 5985 ACTCACCGGTATCATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 6044  
 |||||  
 Db 377 ACTCACCGGTATCATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 436  
 |||||  
 QY 6045 TGAACCCCTCGTACCCCATACTCGCTTGGTAAAGCTATTAAATACCAACCTCACTGGGCT 6104  
 |||||  
 Db 437 TGAACCCCTCGTACCCCATACTCGCTTGGTAAAGCTATTAAATACCAACCTCACTGGGCT 496  
 |||||  
 QY 6105 CCATGAGGTCTCGGCCCAAAACCTCTACTAACTGTGTGGATATGCGCT 6149  
 |||||  
 Db 497 CCATGAGGTCTCGGCCCAAAACCTCTACTAACTGTGTGGATATGCGCT 541  
 |||||

RESULT 23  
 BX368078/c  
 LOCUS BX368078 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CSODI051YM13 3-PRIME, mRNA sequence.  
 ACCESSION BX368078  
 VERSION BX368078.2 GI:46572908  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 846)  
 AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On May 8, 2003 this sequence version replaced gi:30447710.

Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 4215.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0AU0122D05\_U01092\_l&c=4215.r.

FEATURES Location/Qualifiers  
 source 1..846  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODI051YM13"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 4.7%; Score 360; DB 4; Length 846;  
 Best Local Similarity 99.6%; Pred. No. 4e-189;  
 Matches 460; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 6260 TGGAAATAACCATACCTCAAAACCTCAGCTGTGTAAAAATTTAGCAATACTACATACAA 6319  
 |||||  
 Db 629 TGGAAATAACCATACCTCAAAACCTCAGCTGTGTAAAAATTTAGCAATACTACATACAA 570  
 |||||



```
QY 6320 CCACTCCCAATGCATCAGTGGTAACTCTCTCCACACAAAATAGTCTGCTACCTTCAG 6379
|
|
|
Db 569 CCAACTCCCAATGCATCAGTGGTAACTCTCTCCACACAAAATAGTCTGCTACCTTCAG 510
|
|
|
QY 6380 GAATATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTA 6439
|
|
|
Db 509 GAATATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTA 450
|
|
|
QY 6440 TGTGCTTCCTCTCATTTCTTAGTGGCCCCYATGRCATCTACATGAACAAGATTTATACA 6499
|
|
|
Db 449 TGTGCTTCCTCTCATTTCTTAGTGGCCCCYATGRCATCTACATGAACAAGATTTATACA 390
|
|
|
QY 6500 GTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTTGTATAGGAG 6559
|
|
|
Db 389 GTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTTGTATAGGAG 330
|
|
|
QY 6560 CAGGAGTGTAGTGGTACCTAGTGGTATGGCGGTATCACAACTCTCTACTCAGTTCT 6619
|
|
|
Db 329 CAGGAGTGTAGTGGTACCTAGTGGTATGGCGGTATCACAACTCTCTACTCAGTTCT 270
|
|
|
QY 6620 ACTACAAATATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGGACTCCCTGGTCA 6679
|
|
|
Db 269 ACTACAAATATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGGACTCCCTGGTCA 210
|
|
|
QY 6680 CTTTGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCCTTC 6721
|
|
|
Db 209 CTTTGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCCTTC 168
|
|
|
```

```
RESULT 24
DA845887
LOCUS
DEFINITION
  DA845887 PLAC66 Homo sapiens cDNA clone PLAC66014209 5', mRNA
  sequence.
ACCESSION
  DA845887
VERSION
  DA845887.1 GI:81473746
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1. (bases 1 to 547)
  Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
  Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
  Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
  Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
  Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
  Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
  Diversification of Transcriptional Modulation: Large-scale
  Identification and Characterization of Putative Alternative
  Promoters of Human Genes
  Genome Res. 16 (1), 55-65 (2006)
  16344560
  Contact: Takao Isogai
  FLJ Project (HRI Team)
  Helix Research Institute
  2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
  Tel: 81-438-52-3975
  Fax: 81-438-52-3986
  Email: flj-cdna@nifty.com
```

```
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan): cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
```

```
FEATURES
  source
  1..547
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
```

```
/clone="PLAC66014209"
/tissue_type="placenta"
/clone_lib="PLAC66"
/note="Vector: pME18SFL3"
```

```
ORIGIN
Query Match 4.7%; Score 358; DB 9; Length 547;
Best Local Similarity 99.4%; Pred. No. 5.1e-188;
Matches 508; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5748 TAAGGGAACCCCCACCTTCTACCTGCCACACCCATATGCCCGCAACTGCTATCACTCTGC 5807
|
|
|
Db 35 TAAGGGAACCCCCACCTTCTACCTGCCACACCCATATGCCCGCAACTGCTATCACTCTGC 94
|
|
|
QY 5808 CACTCTTTGATGATGCAATCAATCACTATTGAGCAGGAAAATGATTAATCTAGTTG 5867
|
|
|
Db 95 CACTCTTTGATGATGCAATCAATCACTATTGAGCAGGAAAATGATTAATCTAGTTG 154
|
|
|
QY 5868 TCTTGAGGACCTGGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTGA 5927
|
|
|
Db 155 TCTTGAGGACCTGGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTGA 214
|
|
|
QY 5928 TGGGGTGGAGTTCAGAGATCAGCAAGAGAAAACATGTAAPAGAGTAGTAATCTCCCACT 5987
|
|
|
Db 215 TGGGGTGGAGTTCAGAGATCAGCAAGAGAAAACATGTAAPAGAGTAGTAATCTCCCACT 274
|
|
|
QY 5988 CACCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGA 6047
|
|
|
Db 275 CACCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGA 334
|
|
|
QY 6048 AACCTTCCTGATCCTACTGCTGTGAAGCTTATTAATACACCTCACTGGGTCCA 6107
|
|
|
Db 335 AACCTTCCTGATCCTACTGCTGTGAAGCTTATTAATACACCTCACTGGGTCCA 394
|
|
|
QY 6108 TGAGGTCTCGGCCCAAAACCTTACTAACTGTTGGATATGCTCCCTGAACTTCAGGC 6167
|
|
|
Db 395 CGAGGTCTCGGCCCAAAACCTTACTAACTGTTGGATATGCTCCCTGAACTTCAGGC 454
|
|
|
QY 6168 ATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTTCAGCAGAGAATAAACCCAC 6227
|
|
|
Db 455 ATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTTCAGCAGAGAATAAACCCAC 514
|
|
|
QY 6228 TTCCGTTTTAGTAGGACCTCTTGTTCCTCAAT 6258
|
|
|
Db 515 TTCCGTTTTAGTAGGACCTCTTGTTCCTCAAT 545
|
|
|
```

```
RESULT 25
DA846543
LOCUS
DEFINITION
  DA846543 PLAC66 Homo sapiens cDNA clone PLAC66015035 5', mRNA
  sequence.
ACCESSION
  DA846543
VERSION
  DA846543.1 GI:81817963
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1. (bases 1 to 571)
  Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
  Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
  Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
  Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
  Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
  Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
  Diversification of Transcriptional Modulation: Large-scale
  Identification and Characterization of Putative Alternative
```

```
Promoters of Human Genes
  Genome Res. 16 (1), 55-65 (2006)
  16344560
  Contact: Takao Isogai
```

```

FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES             source
    source
        1..571
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="PLACE6015035"
        /tissue_type="placenta"
        /clone_lib="PLACE6"
        /note="Vector: pME18SFL3"

ORIGIN
Query Match          4.7%; Score 357; DB 9; Length 571;
Best Local Similarity 99.6%; Pred. No. 1.9e-187;
Matches 457; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGGAATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 113 AGAGTTTCTATGAGGAATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 172

QY 5745 TTSTAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGGAACCTGCTATCACTC 5804
Db 173 TTCTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGGAACCTGCTATCACTC 232

QY 5805 TGCCACTCTTTCATGCAATGCAAACTCATTTATTGGACAGGAAATGATTAATCCTAG 5864
Db 233 TGCCACTCTTTCATGCAATGCAAACTCATTTATTGGACAGGAAATGATTAATCCTAG 292

QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGCTCTGTTGGACTTACTTCACCCCAACTGGTATGTC 5924
Db 293 TTGTCTCTGGAGGACTTGGAGTCACTGCTCTGTTGGACTTACTTCACCCCAACTGGTATGTC 352

QY 5925 TGATGGGGTGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAGTAATCTCCCA 5984
Db 353 TGATGGGGTGAGTTCAAGATCAGGCAAGAGAGAAAACATGTAAAGAGTAATCTCCCA 412

QY 5985 ACTCACCCGGGTACATGGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACA 6044
Db 413 ACTCACCCGGGTACATGGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACA 472

QY 6045 TGAACCCCTCCGTACCCTATCGCTGCTGTAAGCCCTATTATATACCAACCCCTCACTGGGCT 6104
Db 473 TGAACCCCTCCGTACCCTATCGCTGCTGTAAGCCCTATTATATACCAACCCCTCACTGGGCT 532

QY 6105 CCATGAGGTCCTGGGCCAAACCCCTACTAACTGTTGGAT 6143
Db 533 CCATGAGGTCCTGGGCCAAACCCCTACTAACTGTTGGAT 571

RESULT 26
BX389657
LOCUS
DEFINITION BX389657 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1051YM13 5-PRIME, mRNA sequence.
ACCESSION BX389657
VERSION BX389657.2
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.

```

```

1. (bases 1 to 995)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30462931.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.1
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAF027ZB05_AFO2531_2&c=4215.1

FEATURES             source
    source
        1..995
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0D1051YM13"
        /tissue_type="PLACENTA COT 25-NORMALIZED"
        /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT)
        primer. Five prime end enriched, double-strand cDNA was
        digested with Not I and cloned into the Not I and EcoR V
        sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match          4.7%; Score 356; DB 4; Length 995;
Best Local Similarity 99.3%; Pred. No. 7.1e-187;
Matches 556; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5887 ACTGCTCTTGGACTTACTTCAACCAACTGGTATGCTGGGGTGGAGTTCAAGAT 5946
Db 41 ACTGCTCTTGGACTTACTTCAACCAACTGGTATGCTGGGGTGGAGTTCAAGAT 100

QY 5947 CAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCGGGTACATGGCACC 6006
Db 101 CAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCGGGTACATGGCACC 160

QY 6007 TCTAGCCCTCAAGAGACTAGATCTCTCAAAAATACATGAACCCCTCGTACCATACT 6066
Db 161 TCTAGCCCTCAAGAGACTAGATCTCTCAAAAATACATGAACCCCTCGTACCATACT 220

QY 6067 CGCCTGGTAAGCCTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCCCAAAAC 6126
Db 221 CGCCTGGTAAGCCTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCCCAAAAC 280

QY 6127 CCTACTAATCTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCTGTA 6186
Db 281 CCTACTAATCTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCTGTA 340

QY 6187 CCTGAACAAATGGAACAACCTTCAGCAGAGAAATAAACAACAATTCCTAGTAGACCT 6246
Db 341 CCTGAACAAATGGAACAACCTTCAGCAGAGAAATAAACAACAATTCCTAGTAGACCT 400

QY 6247 CTGTGTTTCCCAATSTGGAATAACCCATACCTCAACCTCACCTGTGTAAATTTAGCAAT 6306
Db 401 CTGTGTTTCCCAATCTGGAATAACCCATACCTCAACCTCACCTGTGTAAATTTAGCAAT 460

QY 6307 ACTACATACACAAACCAACTCCCAATGCATCAGGTGGTAACTCTCCCAACACAAATAGTC 6366
Db 461 ACTACATACACAAACCAACTCCCAATGCATCAGGTGGTAACTCTCCCAACACAAATAGTC 520

QY 6367 TGCCCTACCCCTCAGGAATATTTTGTCTGTGGTACTCTAGCCCTATCGTTTGTGATGCG 6426
Db 521 TGCCCTACCCCTCAGGAATATTTTGTCTGTGGTACTCTAGCCCTATCGTTTGTGATGCG 580

QY 6427 TCTTCAGAAATCTATGTGCTT 6446

```

Db	581	TCTTCAGAAATCTATGTGCTT	600
RESULT 27			
LOCUS	BX430050/c	658 bp	mRNA linear EST 04-MAY-2004
DEFINITION	BX430050 Homo sapiens PLACENTA clone CS0DE012YJ24		
ACCESSION	BX430050		
VERSION	BX430050.2	GI:47003301	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
TITLE	1 (bases 1 to 658)		
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization On May 15, 2003 this sequence version replaced gi:30776872. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r For more information about this cluster, see <a href="http://www.genoscope.cns.fr/cdna?s=CS0BAB002D11_CS00181_1&amp;c=4215.r">http://www.genoscope.cns.fr/cdna?s=CS0BAB002D11_CS00181_1&amp;c=4215.r</a>		
FEATURES	source		
	Location/Qualifiers		
	1..658		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DB012XJ24"		
	/tissue_type="PLACENTA"		
	/clone_lib="Homo sapiens PLACENTA"		
	/note="vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."		
ORIGIN			
	Query Match	4.7%; Score 353; DB 4; Length 658;	
	Best Local Similarity	99.3%; Pred. No. 3.3e-185;	
	Matches	553; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Qy	6165	GCCATATGTTTCAATCCTGTACTGAACAATGGAACTTCAGCACAGAATAAACAC	6224
Db	577	GCCATATGTTTCAATCCTGTACTGAACAATGGAACTTCAGCACAGAATAAACAC	518
Qy	6225	CACCTCGTTTTAGTAGACCTCTTGTTCCCATSTGGAATACCCTACCTCAAACCT	6284
Db	517	CACCTCGTTTTAGTAGACCTCTTGTTCCCATCTGGAATAACCCATACCTCAAACCT	458
Qy	6285	CACCTGTGTAAAATTAGCAATACTACATACACCAACCTCCCAGATGCAGGTGGGT	6344
Db	457	CACCTGTGTAAAATTAGCAATACTACATACACCAACCTCCCAGATGCAGGTGGGT	398
Qy	6345	AACCTCTCCACACAAATAGTCTGCCCTACCTCAGGAATATTTTTGTCTGTGGTACCTC	6404
Db	397	AACCTCTCCACACAAATAGTCTGCCCTACCTCAGGAATATTTTTGTCTGTGGTACCTC	338
Qy	6405	AGCCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTTCTCTCTATTCTAGTGC	6464
Db	337	AGCCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTTCTCTCTATTCTAGTGC	278

```
QY 5866 TGTCTCGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCCACCAACTGGTATGTCT 5925
Db 61 TGTCTCGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCCACCAACTGGTATGTCT 120
QY 5926 GATGGGGGTGGAGTTCAAGATCAGGCAGAGAAAAACATGTAAAAAGAGTATCTCCCAA 5985
Db 121 GATGGGGGTGGAGTTCAAGATCAGGCAGAGAAAAACATGTAAAAAGAGTATCTCCCAA 180
QY 5986 CTCACSSGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACAT 6045
Db 181 CTCACSSGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACAT 240
QY 6046 GAAACCCCTCGTACCCTACTCGCTGTAGAGCTTATTAATACACCCCTCTAGGGTTC 6105
Db 241 GAAACCCCTCGTACCCTACTCGCTGTAGAGCTTATTAATACACCCCTCTAGGGTTC 300
QY 6106 CATGAGGTCTCGGCCCAAAACCTTACTTACTGTGGATGCTCCCTCGAACTTCARG 6165
Db 301 CATGAGGTCTCGGCCCAAAACCTTACTTACTGTGGATGCTCCCTCGAACTTCARG 360
QY 6166 CCATATGTTTCAATCCCTGTACTGAAACAATGAAACAACCTTCAGCACAGAAATAAACACC 6225
Db 361 CCATATGTTTCAATCCCTGTACTGAAACAATGAAACAACCTTCAGCACAGAAATAAACACC 420
QY 6226 ACTTCGGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTTC 6285
Db 421 ACTTCGGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTTC 480
QY 6286 ACCTGTGTAATAATTTAGCAATACTACATACACAACCACTCCCAATGCATCAGGTGGGTA 6345
Db 481 ACCTGTGTAATAATTTAGCAATACTACATACACAACCACTCCCAATGCATCAGGTGGGTA 540
QY 6346 ACTCCTCC 6353
Db 541 ACTCCTCC 548
```

```
RESULT 29
CR735592
LOCUS
DEFINITION
CR735592 NIH_MGC 21 Homo sapiens cDNA clone IMAGE958M21530 ;
IMAGE:3840572 5', mRNA sequence.
ACCESSION
CR735592
VERSION
CR735592.1 GI:51584854
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 554)
Ebert,L., Heil,O., Hennig,S., Korn,B., Neubert,P., Partsch,E.,
Peters,M., Radelof,U. and Schneider,D.
I.M.A.G.E. cDNA Clone Collection
Unpublished (2004)
Contact: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; IMAGE958M21530.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Contact: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.
FEATURES
Location/Qualifiers
1..554
source
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE958M21530 ; IMAGE:3840572"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 21"
/note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dr priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

Query Match	4.5%;	Score 338;	DB 8;	Length 554;
Best Local Similarity	99.5%;	Pred. No. 8.8e-177;		
Matches 438;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY 6282 CCTCACCTGTGTAAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTG 6341
Db 38 CCTCACCTGTGTAAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTG 97
QY 6342 GGTAACTCTCCACACAAAATAGTCTGCCTACCTCAGCAATATTTTGTCTGTGGTAC 6401
Db 98 GGTAACTCTCCACACAAAATAGTCTGCCTACCTCAGCAATATTTTGTCTGTGGTAC 157
QY 6402 CTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAACTATGTCTCTCTCATCTTAGT 6461
Db 158 CTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAACTATGTCTCTCTCATCTTAGT 217
QY 6462 GCCCCYATGCCATCTACACTGAAACAAGATTTATACAGTTATGTATCTAAGCCCG 6521
Db 218 GCCCCCTATGACCATCTACACTGAAACAAGATTTATACAGTTATGTATCTAAGCCCG 277
QY 6522 CAACAAAAGATGACCCATTCTCTTTGTTATAGGACGAGGCTAGGTGCACCTAGG 6581
Db 278 CAACAAAAGATGACCCATTCTCTTTGTTATAGGACGAGGCTAGGTGCACCTAGG 337
QY 6582 TACTGGCAATGGCGGTATCACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAACT 6641
Db 338 TACTGGCAATGGCGGTATCACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAACT 397
QY 6642 AAATGGGACATGGAAACGGGTCGCGACTCCCTGGTCTACCTTGCAGATCAACTTAATCTC 6701
Db 398 AAATGGGACATGGAAACGGGTCGCGACTCCCTGGTCTACCTTGCAGATCAACTTAATCTC 457
QY 6702 CCTAGCAGCAGTAGTCCTTC 6721
Db 458 CCTAGCAGCAGTAGTCCTTC 477
```

RESULT 30  
BX409328  
LOCUS  
DEFINITION  
BX409328 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013120  
5-PRIME, mRNA sequence.  
ACCESSION  
BX409328  
VERSION  
BX409328.2 GI:46932867  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
1 (bases 1 to 924)  
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 13, 2003 this sequence version replaced gi:30652931.  
Contact: Genoscope

Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 4215.r  
For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?s=CS0BAE0122F02\\_AE01071\\_l&c=4215.r](http://www.genoscope.cns.fr/cdna?s=CS0BAE0122F02_AE01071_l&c=4215.r)

## FEATURES

source Location/Qualifiers  
1. .924  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DB013Y120"  
/tissue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

## ORIGIN

Query Match 4.4%; Score 336; DB 4; Length 924;  
Best Local Similarity 99.5%; Pred. No. 1.2e-175;  
Matches 436; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5685 AGAGTTTCTATGAGAAATGAGGTCCTCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744  
Db 155 AGAGTTTCTATGAGAAATGAGGTCCTCCGGAATATTGATGCCCATCGTATAGGAGTCT 214  
QY 5745 TTSTAAGGGAAACCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804  
Db 215 TTCTAAGGGAAACCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 274  
QY 5805 TGCACCTCTTTGCATGATGCAATACTCATTTATGGACAGGAAATGATTAATCCTAG 5864  
Db 275 TGCACCTCTTTGCATGATGCAATACTCATTTATGGACAGGAAATGATTAATCCTAG 334  
QY 5865 TTGTCTTGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTAGTC 5924  
Db 335 TTGTCTTGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTAGTC 394  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAAATCTCCCA 5984  
Db 395 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAAATCTCCCA 454  
QY 5985 ACTCACCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 6044  
Db 455 ACTCACCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 514  
QY 6045 TGAACCCCTCCGTACCCATCTCGCTGGTAAAGCTATTTAATACACCCCTCACTGGGCT 6104  
Db 515 TGAACCCCTCCGTACCCATCTCGCTGGTAAAGCTATTTAATACACCCCTCACTGGGCT 574  
QY 6105 CCATGAGGTCTCGGCCCA 6122  
Db 575 CCATGAGGTCTCGGCCCA 592

Search completed: August 2, 2006, 15:18:55  
Job time : 21283 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM nucleic - nucleic search, using sw model  
 Run on: August 2, 2006, 09:10:45 ; Search time 436 Seconds  
 (without alignments)  
 9710.509 Million cell updates/sec

Title: US-10-717-580-11  
 Perfect score: 7582  
 Sequence: 1 caacaatcggtataaacc.....tattaatcttgcartcgt 7582

Scoring table: OLIGO\_NUC  
 Gapop 60.0 , Gapext 60.0

Searched: 1452992 seqs, 279199553 residues

Word size : 6

Total number of hits satisfying chosen parameters: 2903664

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : Published Applications NA New.\*  
 1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
 2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
 3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
 4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
 5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
 6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
 7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
 8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	0.4	44063	6 US-10-539-228-718	Sequence 718, App
2	28	0.4	138627	6 US-10-540-898-159	Sequence 159, App
3	22	0.3	35	6 US-10-544-059-33	Sequence 33, Appl
4	22	0.3	36	6 US-10-544-059-41	Sequence 41, Appl
5	22	0.3	39	6 US-10-544-059-30	Sequence 30, Appl
6	21	0.3	36	6 US-10-544-059-37	Sequence 37, Appl
7	21	0.3	40	6 US-10-544-059-32	Sequence 32, Appl
8	21	0.3	170489	6 US-10-539-228-380	Sequence 380, App
9	20	0.3	20	6 US-10-554-711-771	Sequence 771, App
10	20	0.3	20	6 US-10-554-711-775	Sequence 775, App
11	20	0.3	38	6 US-10-544-059-36	Sequence 36, Appl
12	20	0.3	50	6 US-10-544-711-192	Sequence 192, App
13	20	0.3	1507	7 US-11-056-355B-55306	Sequence 55306, A
14	20	0.3	256190	6 US-10-539-228-320	Sequence 320, App
15	19	0.3	35	6 US-10-544-059-40	Sequence 40, Appl
16	19	0.3	35	6 US-10-544-059-42	Sequence 42, Appl
17	19	0.3	35	6 US-10-544-059-44	Sequence 44, Appl
18	19	0.3	1877	6 US-10-449-902-25572	Sequence 25572, A
19	19	0.3	2047	6 US-10-516-094-14	Sequence 14, Appl
20	19	0.3	2994	6 US-10-540-898-5	Sequence 5, Appl
21	19	0.3	84476	6 US-10-540-898-4	Sequence 4, Appl
22	19	0.3	225587	6 US-10-540-898-374	Sequence 374, App
23	18	0.2	18	6 US-10-554-711-777	Sequence 777, App
24	18	0.2	33	6 US-10-544-059-31	Sequence 31, Appl
25	18	0.2	33	6 US-10-544-059-39	Sequence 39, Appl

## RESULT 1

US-10-539-228-718

; Sequence 718, Application US/10539228

## ALIGNMENTS

26	18	0.2	33	6 US-10-544-059-43	Sequence 43, Appl
27	18	0.2	40	6 US-10-544-059-34	Sequence 34, Appl
28	18	0.2	561	7 US-11-300-928-23	Sequence 23, Appl
29	18	0.2	575	6 US-10-488-619-2900	Sequence 2900, Ap
30	18	0.2	840	7 US-11-218-305-24411	Sequence 24411, A
31	18	0.2	1080	6 US-10-471-571A-4897	Sequence 4897, Ap
32	18	0.2	1095	6 US-10-471-571A-557	Sequence 557, App
33	18	0.2	1099	7 US-11-056-355B-58368	Sequence 58368, A
34	18	0.2	1109	6 US-10-449-902-2644	Sequence 2644, Ap
35	18	0.2	1139	7 US-11-218-305-2912	Sequence 2912, Ap
36	18	0.2	1223	6 US-10-449-902-2033	Sequence 2033, Ap
37	18	0.2	1273	7 US-11-218-305-13693	Sequence 13693, A
38	18	0.2	1285	6 US-10-449-902-26581	Sequence 26581, A
39	18	0.2	1332	6 US-10-449-902-2328	Sequence 2328, Ap
40	18	0.2	1352	6 US-10-449-902-2915	Sequence 2915, Ap
41	18	0.2	1377	6 US-10-449-902-21352	Sequence 21352, A
42	18	0.2	1395	6 US-10-449-902-1867	Sequence 1867, Ap
43	18	0.2	1420	6 US-10-449-902-23075	Sequence 23075, A
44	18	0.2	1462	6 US-10-449-902-947	Sequence 947, App
45	18	0.2	1463	6 US-10-449-902-3077	Sequence 3077, App
46	18	0.2	1480	6 US-10-449-902-16130	Sequence 16130, A
47	18	0.2	1511	7 US-11-218-305-13694	Sequence 13694, A
48	18	0.2	1559	7 US-11-218-305-17084	Sequence 17084, A
49	18	0.2	1572	7 US-11-056-355B-64677	Sequence 64677, A
50	18	0.2	1634	6 US-10-449-902-1162	Sequence 1162, Ap
51	18	0.2	1664	6 US-10-449-902-4521	Sequence 4521, Ap
52	18	0.2	1684	6 US-10-449-902-9631	Sequence 9631, Ap
53	18	0.2	1779	6 US-10-449-902-26973	Sequence 26973, A
54	18	0.2	1822	6 US-10-449-902-8489	Sequence 8489, Ap
55	18	0.2	1872	6 US-10-449-902-2385	Sequence 2385, Ap
56	18	0.2	1896	6 US-10-449-902-1149	Sequence 1149, Ap
57	18	0.2	1935	6 US-10-449-902-9739	Sequence 9739, Ap
58	18	0.2	2132	6 US-10-449-902-17925	Sequence 17925, A
59	18	0.2	2162	6 US-10-449-902-14299	Sequence 14299, A
60	18	0.2	2187	6 US-10-449-902-28272	Sequence 28272, A
61	18	0.2	2369	7 US-11-056-355B-28466	Sequence 28466, A
62	18	0.2	2369	7 US-11-056-355B-33076	Sequence 33076, A
63	18	0.2	2390	6 US-10-449-902-8396	Sequence 8396, Ap
64	18	0.2	2448	7 US-11-056-355B-87454	Sequence 87454, A
65	18	0.2	2527	7 US-11-218-305-9934	Sequence 9934, Ap
66	18	0.2	2589	6 US-10-449-902-6620	Sequence 6620, Ap
67	18	0.2	2597	6 US-10-539-228-269	Sequence 269, App
68	18	0.2	2905	6 US-10-449-902-23130	Sequence 23130, A
69	18	0.2	3145	7 US-11-218-305-9936	Sequence 9936, Ap
70	18	0.2	3614	6 US-10-449-902-13810	Sequence 13810, A
71	18	0.2	4025	6 US-10-449-902-12645	Sequence 12645, A
72	18	0.2	7230	7 US-11-218-305-18941	Sequence 18941, A
73	18	0.2	32847	6 US-10-540-898-744	Sequence 744, App
74	18	0.2	70665	6 US-10-505-928-596	Sequence 596, App
75	18	0.2	107432	6 US-10-540-898-137	Sequence 137, App
76	18	0.2	107745	6 US-10-539-228-268	Sequence 268, App
77	18	0.2	108057	6 US-10-540-898-835	Sequence 835, App
78	18	0.2	180026	6 US-10-540-898-729	Sequence 729, App
79	18	0.2	195102	6 US-10-733-042-1	Sequence 1, Appl
80	18	0.2	257645	7 US-10-540-898-266	Sequence 266, App
81	17	0.2	25	7 US-11-348-413-1043443	Sequence 1043443
82	17	0.2	50	6 US-10-511-937-10	Sequence 10, Appl
83	17	0.2	50	6 US-10-554-711-126	Sequence 126, App
84	17	0.2	240	7 US-11-348-413-12665	Sequence 12665, A
85	17	0.2	494	7 US-11-376-673-29	Sequence 29, Appl
86	17	0.2	501	7 US-11-056-355B-78211	Sequence 78211, A
87	17	0.2	574	6 US-10-488-619-1539	Sequence 1539, Ap
88	17	0.2	611	6 US-10-541-993-66	Sequence 66, Appl
89	17	0.2	611	6 US-10-541-993-69	Sequence 69, Appl
90	17	0.2	723	6 US-10-449-902-7472	Sequence 7472, Ap



; Publication No. US20060154250A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)  
; CURRENT APPLICATION NUMBER: US/10/539,228  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 718  
; LENGTH: 44063  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(44063)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-718

Query Match 0.4%; Score 34; DB 6; Length 44063;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5709 TCCCGGAATATGTATGCCCATCGTATAGGAGT 5742  
|||||  
DB 916 TCCCGGAATATGTATGCCCATCGTATAGGAGT 949  
|||||

## RESULT 2

US-10-540-898-159/c  
; Sequence 159, Application US/10540898  
; Publication No. US20060166213A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 159  
; LENGTH: 138627  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(138627)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-540-898-159

Query Match 0.4%; Score 28; DB 6; Length 138627;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GGGATATAAACCCAGGCATTCGAGCTGG 36  
|||||  
DB 81809 GGGATATAAACCCAGGCATTCGAGCTGG 81782  
|||||

## RESULT 3

US-10-544-059-33  
; Sequence 33, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.

; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-33

Query Match 0.3%; Score 22; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6614 AGTTCTACTACAAACTATCTCA 6635  
|||||  
DB 14 AGTTCTACTACAAACTATCTCA 35  
|||||

## RESULT 4

US-10-544-059-41  
; Sequence 41, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 36  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-41

Query Match 0.3%; Score 22; DB 6; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 GTTTGGCGATCTCTGGTATCTC 1291  
|||||  
DB 15 GTTTGGCGATCTCTGGTATCTC 36  
|||||

## RESULT 5

US-10-544-059-30/c  
; Sequence 30, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 30  
; LENGTH: 39  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-30

Query Match 0.3%; Score 22; DB 6; Length 39;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6135 CTCTGGATATGCTCCCGCTG 6156  
|||||  
Db 39 CTCTGGATATGCTCCCGCTG 18

## RESULT 6

US-10-544-059-37  
; Sequence 37, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: LATERAL SCLEROSIS  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 36  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-37

Query Match 0.3%; Score 21; DB 6; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 963 CAAAGGAGATACACAAAGG 983  
|||||  
Db 16 CAAAGGAGATACACAAAGG 36

## RESULT 7

US-10-544-059-32/C  
; Sequence 32, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: LATERAL SCLEROSIS  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32  
; LENGTH: 40  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-32

Query Match 0.3%; Score 21; DB 6; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6512 CTAAGCCCGCAACAAAGAG 6532  
|||||  
Db 40 CTAAGCCCGCAACAAAGAG 20

## RESULT 8

US-10-539-228-380  
; Sequence 380, Application US/10539228  
; Publication No. US20060154250A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)  
; CURRENT APPLICATION NUMBER: US/10/539,228  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 380  
; LENGTH: 170489  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(170489)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-380

Query Match 0.3%; Score 21; DB 6; Length 170489;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5994 GGTACATGGCACCTCTAGCCC 6014  
|||||  
Db 167429 GGTACATGGCACCTCTAGCCC 167449

## RESULT 9

US-10-554-711-771  
; Sequence 771, Application US/10554711  
; Publication No. US20060115806A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Georgia Research Foundation  
; APPLICANT: McDonald, John F.  
; TITLE OF INVENTION: Global Analysis of Transposable Elements  
; FILE REFERENCE: as Molecular Markers of Cancer  
; CURRENT APPLICATION NUMBER: US/10/554,711  
; CURRENT FILING DATE: 2005-10-28  
; PRIOR APPLICATION NUMBER: 60/466,798  
; PRIOR FILING DATE: 2003-04-29  
; NUMBER OF SEQ ID NOS: 778  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 771  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/note =  
US-10-554-711-771

Query Match 0.3%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 CCACCACCTGCTGTTGCCAC 190  
|||||  
Db 1 CCACCACCTGCTGTTGCCAC 20

```
RESULT 10
US-10-554-711-775
; Sequence 775, Application US/10554711
; Publication No. US20060115806A1
; GENERAL INFORMATION:
; APPLICANT: University of Georgia Research Foundation
; APPLICANT: McDonald, John F.
; TITLE OF INVENTION: Global Analysis of Transposable Elements
; FILE OF INVENTION: as Molecular Markers of Cancer
; FILE REFERENCE: 21099.0075P1
; CURRENT APPLICATION NUMBER: US/10/554,711
; CURRENT FILING DATE: 2005-10-28
; PRIOR APPLICATION NUMBER: 60/466,798
; PRIOR FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 778
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: Synthetic Construct
US-10-554-711-775

Query Match      0.3%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6590 TTGGCGGTATCATCAACCTCT 6609
Db 1 TTGGCGGTATCATCAACCTCT 20

RESULT 11
US-10-544-059-36/c
; Sequence 36, Application US/10544059
; Publication No. US20060160087A1
; GENERAL INFORMATION:
; APPLICANT: SLIL Biomedical Corporation
; APPLICANT: McGrath, Michael
; APPLICANT: Hadlock, Kenneth G.
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC
; FILE REFERENCE: 376462002040
; CURRENT APPLICATION NUMBER: US/10/544,059
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: US 60/444,224
; PRIOR FILING DATE: 2003-01-31
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Herv-W
US-10-544-059-36

Query Match      0.3%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1005 GCCAATATTCCTCCCAATTATG 1024
Db 38 GCCAATATTCCTCCCAATTATG 19

RESULT 12
US-10-554-711-192
; Sequence 192, Application US/10554711
; Publication No. US20060115806A1
; GENERAL INFORMATION:
; APPLICANT: University of Georgia Research Foundation
```

```
; APPLICANT: McDonald, John F.
; TITLE OF INVENTION: Global Analysis of Transposable Elements
; FILE OF INVENTION: as Molecular Markers of Cancer
; FILE REFERENCE: 21099.0075P1
; CURRENT APPLICATION NUMBER: US/10/554,711
; CURRENT FILING DATE: 2005-10-28
; PRIOR APPLICATION NUMBER: 60/466,798
; PRIOR FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 778
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: Synthetic Construct
US-10-554-711-192

Query Match      0.3%; Score 20; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 484 GCTTGCCACCATCTTGGGAAG 503
Db 29 GCTTGCCACCATCTTGGGAAG 48

RESULT 13
US-11-056-355B-55306
; Sequence 55306, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590FUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 55306
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1507)
; OTHER INFORMATION: Ceres Seq. ID no. 12607536
US-11-056-355B-55306

Query Match      0.3%; Score 20; DB 7; Length 1507;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6439 ATGTGTTCTCTCTCATCTT 6458
Db 736 ATGTGTTCTCTCTCATCTT 755

RESULT 14
US-10-539-228-320/c
; Sequence 320, Application US/10539228
; Publication No. US20060154250A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)
; CURRENT APPLICATION NUMBER: US/10/539,228
; CURRENT FILING DATE: 2005-06-17
```

; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 320  
; LENGTH: 256190  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(256190)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-320

Query Match 0.3%; Score 20; DB 6; Length 256190;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1502 CACAGGGGAAGGAAGAAA 1521  
|||||  
DB 88993 CACAGGGGAAGGAAGAAA 88974

## RESULT 15

US-10-544-059-40/c  
; Sequence 40, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: LATERAL SCLEROSIS  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-40

Query Match 0.3%; Score 19; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 GTTTTCAAGGGTTAGGAC 1174  
|||||  
DB 35 GTTTTCAAGGGTTAGGAC 17

## RESULT 16

US-10-544-059-42/c  
; Sequence 42, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: LATERAL SCLEROSIS  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42

; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-42

Query Match 0.3%; Score 19; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1520 AATCCTACTGCCTTCTGG 1538  
|||||  
DB 35 AATCCTACTGCCTTCTGG 17

## RESULT 17

US-10-544-059-44/c  
; Sequence 44, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: LATERAL SCLEROSIS  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-44

Query Match 0.3%; Score 19; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1622 ATCACTCAGTCAGCTGCAG 1640  
|||||  
DB 35 ATCACTCAGTCAGCTGCAG 17

## RESULT 18

US-10-449-902-25572  
; Sequence 25572, Application US/10449902  
; Publication No. US20060123505A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agrobiological Sciences.  
; APPLICANT: Bio-oriented Technology Research Advancement Institution.  
; APPLICANT: The Institute of Physical and Chemical Research.  
; APPLICANT: Foundation for Advancement of International Science.  
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF  
; FILE REFERENCE: MOA-A0205V1-US  
; CURRENT APPLICATION NUMBER: US/10/449,902  
; CURRENT FILING DATE: 2003-05-29  
; PRIOR APPLICATION NUMBER: JP 2002-203269  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: JP 2002-383870  
; PRIOR FILING DATE: 2002-12-11  
; NUMBER OF SEQ ID NOS: 56791  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25572  
; LENGTH: 1877  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: AK101014  
; DATABASE ENTRY DATE: 2002-08-28  
US-10-449-902-25572

Query Match 0.3%; Score 19; DB 6; Length 1877;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1533 TTCTGGAGAGACTAAGGGA 1551

DB 419 TTCTGGAGAGACTAAGGGA 437

## RESULT 19

US-10-516-094-14/c  
; Sequence 14, Application US/10516094  
; Publication No. US20060141457A1  
; GENERAL INFORMATION:  
; APPLICANT: Dahlqvist, Anders  
; APPLICANT: Ghosal, Alakesh  
; APPLICANT: Lindqvist, Ylva  
; APPLICANT: Banas, Antoni  
; TITLE OF INVENTION: NEW IMPROVED ACYLTRANSFERASE  
; FILE REFERENCE: 5822.317USWO  
; CURRENT APPLICATION NUMBER: US/10/516,094  
; CURRENT FILING DATE: 2004-11-24  
; PRIOR APPLICATION NUMBER: PCT/SE03/00870  
; PRIOR FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: SE 0201581-6  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: SE 0300142-7  
; PRIOR FILING DATE: 2003-01-20  
; PRIOR APPLICATION NUMBER: US 60/383,889  
; PRIOR FILING DATE: 2002-05-29  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 14  
; LENGTH: 2047  
; TYPE: DNA  
; ORGANISM: Aspergillus nidulans  
US-10-516-094-14

Query Match 0.3%; Score 19; DB 6; Length 2047;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6442 TGCTTCTCTCATCTTAG 6460

DB 396 TGCTTCTCTCATCTTAG 378

## RESULT 20

US-10-540-898-5/c  
; Sequence 5, Application US/10540898  
; Publication No. US20060166213A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2994  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-540-898-5

Query Match 0.3%; Score 19; DB 6; Length 2994;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1976 TGCCCCAGGGGACAAAGGT 1994

DB 1594 TGCCCCAGGGGACAAAGGT 1576

## RESULT 21

US-10-540-898-4/c  
; Sequence 4, Application US/10540898  
; Publication No. US20060166213A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 84476  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-540-898-4

Query Match 0.3%; Score 19; DB 6; Length 84476;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1976 TGCCCCAGGGGACAAAGGT 1994

DB 73076 TGCCCCAGGGGACAAAGGT 73058

## RESULT 22

US-10-540-898-374/c  
; Sequence 374, Application US/10540898  
; Publication No. US20060166213A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 374  
; LENGTH: 225587  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(225587)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-540-898-374

Query Match 0.3%; Score 19; DB 6; Length 225587;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6992 TCCTTGTAACTTTGTCTC 7010

DB 218645 TCCTTGTAACTTTGTCTC 218627

## RESULT 23

US-10-554-711-777/c  
; Sequence 777, Application US/10554711  
; Publication No. US20060115806A1

; GENERAL INFORMATION:  
; APPLICANT: University of Georgia Research Foundation  
; APPLICANT: McDonald, John F.  
; TITLE OF INVENTION: Global Analysis of Transposable Elements  
; FILE REFERENCE: 21099.0075P1  
; CURRENT APPLICATION NUMBER: US/10/554,711  
; PRIOR FILING DATE: 2005-10-28  
; PRIOR APPLICATION NUMBER: 60/466,798  
; PRIOR FILING DATE: 2003-04-29  
; NUMBER OF SEQ ID NOS: 778  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 777  
; LENGTH: 18  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/note =  
; OTHER INFORMATION: Synthetic Construct  
US-10-554-711-777

Query Match 0.2%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6807 TCAATCCGGAATCGTCAC 6824  
Db 18 TCAATCCGGAATCGTCAC 1

## RESULT 24

US-10-544-059-31  
; Sequence 31, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31  
; LENGTH: 33  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-31

Query Match 0.2%; Score 18; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5935 GGAGTTCAAGATCAGGCA 5952  
Db 16 GGAGTTCAAGATCAGGCA 33

## RESULT 25

US-10-544-059-39  
; Sequence 39, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040

; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 33  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-39

Query Match 0.2%; Score 18; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 TTATGCCCTACAGGAAGC 883  
Db 16 TTATGCCCTACAGGAAGC 33

## RESULT 26

US-10-544-059-43  
; Sequence 43, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 33  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-43

Query Match 0.2%; Score 18; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1486 ATGATGTCCACCAACA 1503  
Db 16 ATGATGTCCACCAACA 33

## RESULT 27

US-10-544-059-34/c  
; Sequence 34, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 40  
; TYPE: DNA

```
; ORGANISM: Herv-W
US-10-544-059-34

Query Match          0.2%; Score 18; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6890 CCTGGGGCTCTCTCAGCC 6907
Db 40 CCTGGGGCTCTCTCAGCC 23

RESULT 28
US-11-300-928-23
; Sequence 23, Application US/11300928
; Publication No. US20060166277A1
; GENERAL INFORMATION:
; APPLICANT: Karumanchi, S. Ananth
; APPLICANT: Sukhatme, Vikas P.
; TITLE OF INVENTION: Nucleic Acids and Polypeptides Useful For Diagnosing and Treating
; FILE REFERENCE: 01948/108002
; CURRENT APPLICATION NUMBER: US/11/300,928
; CURRENT FILING DATE: 2005-12-15
; PRIOR APPLICATION NUMBER: US 60/636,275
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-300-928-23

Query Match          0.2%; Score 18; DB 7; Length 561;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 TCCCAAGCAGTGGAGGA 1047
Db 342 TCCCAAGCAGTGGAGGA 359

RESULT 29
US-10-488-619-2900
; Sequence 2900, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P. C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2900
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2900

Query Match          0.2%; Score 18; DB 6; Length 575;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 ATGGGAGCTGTTTCATG 89
Db 430 ATGGGAGCTGTTTCATG 447

RESULT 30
US-11-218-305-24411
```

```
; Sequence 24411, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McIaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (536660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24411
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (777)..(777)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-24411

Query Match          0.2%; Score 18; DB 7; Length 840;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5687 AGTTTCTATGGAGATGC 5704
Db 136 AGTTTCTATGGAGATGC 153

Search completed: August 2, 2006, 09:18:12
Job time : 439 secs
```



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:46 ; Search time 7477 Seconds  
(without alignments)  
12460.191 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7582  
Sequence: 1 caacaatcgggatataaacc.....tattaaatcttgcaactgcr 7582

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 18892170 seqs, 6143817638 residues

Word size : 6

Total number of hits satisfying chosen parameters: 37734928

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 90 summaries

```

Database : Published Applications NA Main:*
1: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US07_PUBCOMB.seq:*
2: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US08_PUBCOMB.seq:*
3: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US09A_PUBCOMB.seq:*
4: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US09B_PUBCOMB.seq:*
5: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US09C_PUBCOMB.seq:*
6: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US10A_PUBCOMB.seq:*
7: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US10B_PUBCOMB.seq:*
8: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US10C_PUBCOMB.seq:*
9: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US10D_PUBCOMB.seq:*
10: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US10E_PUBCOMB.seq:*
11: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US10F_PUBCOMB.seq:*
12: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US10G_PUBCOMB.seq:*
13: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US11A_PUBCOMB.seq:*
14: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US11B_PUBCOMB.seq:*
15: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US11C_PUBCOMB.seq:*
16: /EMC Celerra_SIDS3/ptodata/2/pubnpa/US11D_PUBCOMB.seq:*

```

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	7582	100.0	7582	8	US-10-632-793-30	Sequence 30, Appl
2	7582	100.0	7582	9	US-10-717-580-11	Sequence 11, Appl
3	979	12.9	7466	10	US-10-450-763-4430	Sequence 4430, Ap
4	888	11.7	2372	8	US-10-632-793-29	Sequence 29, Appl
5	888	11.7	2372	9	US-10-717-580-10	Sequence 10, Appl
6	779	10.3	2599	13	US-11-028-539-1	Sequence 1, Appl
7	779	10.3	8523	3	US-09-854-867-21	Sequence 21, Appl
8	779	10.3	8523	10	US-10-786-970A-21	Sequence 21, Appl
9	779	10.3	10499	13	US-11-028-539-3	Sequence 3, Appl
10	779	10.3	56093	3	US-09-873-367C-81	Sequence 81, Appl
11	779	10.3	56093	10	US-10-643-641A-81	Sequence 81, Appl
12	750	9.9	6394	10	US-10-450-763-20014	Sequence 20014, A
13	731	9.6	1617	8	US-10-133-036-3	Sequence 3, Appl
14	731	9.6	1948	8	US-10-632-793-24	Sequence 24, Appl
15	731	9.6	1948	9	US-10-717-580-5	Sequence 5, Appl
16	731	9.6	2055	13	US-11-028-539-22	Sequence 22, Appl
17	731	9.6	2782	8	US-10-632-793-26	Sequence 26, Appl

ALIGNMENTS

```

RESULT 1
US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauca
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-30

```

Query Match 100.0%; Score 7582; DB 8; Length 7582;

		Best Local Similarity 100.0%; Pred. No. 0;				Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CAACAATCGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC	60						
DB	1	CAACAATCGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC	60						
QY	61	CTTCCCTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCACCTGCA	120						
DB	61	CTTCCCTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCACCTGCA	120						
QY	121	CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCCTGC	180						
DB	121	CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCCTGC	180						
QY	181	TGTTTGGCACACCCGCANACCTGCCGTGACTCCCATCCCTCTGGATCCTGCAGGCTGC	240						
DB	181	TGTTTGGCACACCCGCANACCTGCCGTGACTCCCATCCCTCTGGATCCTGCAGGCTGC	240						
QY	241	CGCTGTGCTCTGATCCAGCGARGGGCCCATTTGGGCTCCCAATTTGGCTAAAGGCTTGC	300						
DB	241	CGCTGTGCTCTGATCCAGCGARGGGCCCATTTGGGCTCCCAATTTGGCTAAAGGCTTGC	300						
QY	301	CATTGTNCTGACGGCTAAGTGTGCTGGTGTGTTCTAATTGAGCTGAACACTANTCACT	360						
DB	301	CATTGTNCTGACGGCTAAGTGTGCTGGTGTGTTCTAATTGAGCTGAACACTANTCACT	360						
QY	361	GGGTTCCATGTTCTCTTCTGTGACCCACCGGCTTCTAATAKAACCTATAACACTTTACCA	420						
DB	361	GGGTTCCATGTTCTCTTCTGTGACCCACCGGCTTCTAATAKAACCTATAACACTTTACCA	420						
QY	421	TGGCCCAAGATTCATTCCTTGGAAATCGTGAGGSCAAAGAACTCCAGGTGAGAGAAATAC	480						
DB	421	TGGCCCAAGATTCATTCCTTGGAAATCGTGAGGSCAAAGAACTCCAGGTGAGAGAAATAC	480						
QY	481	GARGCTTGCCACCATCTTGGAAAGCGGCTGTACRTCTTGGAAAGTGTGTACACCACTTC	540						
DB	481	GARGCTTGCCACCATCTTGGAAAGCGGCTGTACRTCTTGGAAAGTGTGTACACCACTTC	540						
QY	541	TGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCAACGAGGACATCC	600						
DB	541	TGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCAACGAGGACATCC	600						
QY	601	MAAGTATGGGAAACGTTCCCGCAAGCAAAAACGCCCTTAAGACGCTATTCTGGARAT	660						
DB	601	MAAGTATGGGAAACGTTCCCGCAAGCAAAAACGCCCTTAAGACGCTATTCTGGARAT	660						
QY	661	TGGGAMCAATTTGACCTTCAGACACTAAGAAGAAACGACTTATATTCTTCTGCAAGTGC	720						
DB	661	TGGGAMCAATTTGACCTTCAGACACTAAGAAGAAACGACTTATATTCTTCTGCAAGTGC	720						
QY	721	GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACTCTTTTG	780						
DB	721	GCCTGGCACTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACTCTTTTG	780						
QY	781	TAGAAAAGGCAAAATGGAGTGAAGTCCATAAGTACAAAATTTCTTTTCAATTAAGACAA	840						
DB	781	TAGAAAAGGCAAAATGGAGTGAAGTCCATAAGTACAAAATTTCTTTTCAATTAAGACAA	840						
QY	841	CTCACAATTATGTAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC	900						
DB	841	CTCACAATTATGTAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC	900						
QY	901	CTATCCAGCATCCCGACTCCTTCCCCAMTATAAGGACCCGCCCTTCAACCCAAATGG	960						
DB	901	CTATCCAGCATCCCGACTCCTTCCCCAMTATAAGGACCCGCCCTTCAACCCAAATGG	960						
QY	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAAACCAAGAGTGCCTAATATTCGCCAAT	1020						
DB	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAAACCAAGAGTGCCTAATATTCGCCAAT	1020						
QY	1021	TATGACCCCTCCCAAGCAGTGGAGGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCYT	1080						
DB	1021	TATGACCCCTCCCAAGCAGTGGAGGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCYT	1080						

Db	1021	TATGACCCCTCCCAAGCAGTGGGAGGAAGAATTTCGCCCCAGCCAGAGTGCATGTGCTT	1080
QY	1081	TTTTTCTCCAGACTTAAAGCAATAAATAACAGACTTAGGTAAATTTCTCAGATAAYCCT	1140
Db	1081	TTTTTCTCCAGACTTAAAGCAATAAATAACAGACTTAGGTAAATTTCTCAGATAAYCCT	1140
QY	1141	GATGGCTATATGTGTGTTTACAAAGGTTAGGACAATTTCTTGATCTGACATGGAGAGAT	1200
Db	1141	GATGGCTATATGTGTGTTTACAAAGGTTAGGACAATTTCTTGATCTGACATGGAGAGAT	1200
QY	1201	ATATATGCTACTGCTTAATCAGACATAACCCCAAAATGAGAGAAGTGCCACCAATACTGC	1260
Db	1201	ATATATGCTACTGCTTAATCAGACATAACCCCAAAATGAGAGAAGTGCCACCAATACTGC	1260
QY	1261	AGCCTGAGGTGTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA	1320
Db	1261	AGCCTGAGGTGTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACAACA	1320
QY	1321	GAAGGAAGANAATGATTTCCCAAGCAGCAGCAGTTCGCCAGTCTASACCTCATTTG	1380
Db	1321	GAAGGAAGANAATGATTTCCCAAGCAGCAGCAGTTCGCCAGTCTASACCTCATTTG	1380
QY	1381	GGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGC	1440
Db	1381	GGGACACAGAAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGC	1440
QY	1441	TASAGGACTTAAGGAATACTASAGAAATCTAYGAATTACTCAATGATGTCCACATA	1500
Db	1441	TASAGGACTTAAGGAATACTASAGAAATCTAYGAATTACTCAATGATGTCCACATA	1500
QY	1501	ACACAGGGAAGGAAGAAAATCCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAG	1560
Db	1501	ACACAGGGAAGGAAGAAAATCCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAG	1560
QY	1561	GAAAGCGTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT	1620
Db	1561	GAAAGCGTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT	1620
QY	1621	TATCACTCAGTCAGTCGACATTAAGAAAACTTCAAAAGTCTGCCGTAGSCCCGGAG	1680
Db	1621	TATCACTCAGTCAGTCGACATTAAGAAAACTTCAAAAGTCTGCCGTAGSCCCGGAG	1680
QY	1681	CAAACTTGAACACCTATTTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA	1740
Db	1681	CAAACTTGAACACCTATTTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA	1740
QY	1741	GAGCAGCGGGAACAGGACAAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCATGACCTT	1800
Db	1741	GAGCAGCGGGAACAGGACAAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCATGACCTT	1800
QY	1801	CAGGCAAGTGGACTTTTGGAGGCTCTGGAAGGGAAGGCTGGGCAAAATTTGAATGCTTAA	1860
Db	1801	CAGGCAAGTGGACTTTTGGAGGCTCTGGAAGGGAAGGCTGGGCAAAATTTGAATGCTTAA	1860
QY	1861	TAGGGCTTGTCTCAGTGGGTCTACAAGGACACTTTTAAAGATGTGCTCAAGTAGAAG	1920
Db	1861	TAGGGCTTGTCTCAGTGGGTCTACAAGGACACTTTTAAAGATGTGCTCAAGTAGAAG	1920
QY	1921	TAAAGCGCCCTCTGCTCAGTGGCTTTTCAAGGGAATCACTGGAAGGCCACCTGCC	1980
Db	1921	TAAAGCGCCCTCTGCTCAGTGGCTTTTCAAGGGAATCACTGGAAGGCCACCTGCC	1980
QY	1981	CAGGGGCAAAAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCCAGCAGCAGGACTG	2040
Db	1981	CAGGGGCAAAAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCCAGCAGCAGGACTG	2040
QY	2041	AGGGTCCCTGGGCAAGGCGCATGCCATGCCATCACAGCCCTCAGAGCCCTGGGTATGCTT	2100
Db	2041	AGGGTCCCTGGGCAAGGCGCATGCCATGCCATCACAGCCCTCAGAGCCCTGGGTATGCTT	2100
QY	2101	GACCATTTAGGGCCAGAGGTTGTCTCTGGACACTGGTGGGCTTCTTAGTCTTACT	2160
Db	2101	GACCATTTAGGGCCAGAGGTTGTCTCTGGACACTGGTGGGCTTCTTAGTCTTACT	2160
QY	2161	CTTCTGTCCGGACAACCTGTCTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG	2220
Db	2161	CTTCTGTCCGGACAACCTGTCTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG	2220
QY	2221	GGCAGTCACTAGATACCTTTTCCAGGCACTAAGTTATGAATCTGGGAGGCTTTATTTCTT	2280
Db	2221	GGCAGTCACTAGATACCTTTTCCAGGCACTAAGTTATGAATCTGGGAGGCTTTATTTCTT	2280
QY	2281	TCACATGCTTTTCTAAATATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTA	2340
Db	2281	TCACATGCTTTTCTAAATATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTA	2340
QY	2341	GCMAAGCAGGGGCCATTTATACCTGAAATAGGAGAGGAAACAACCCGTTTGTGTGTTCC	2400
Db	2341	GCMAAGCAGGGGCCATTTATACCTGAAATAGGAGAGGAAACAACCCGTTTGTGTGTTCC	2400
QY	2401	CTGTCTTGAAGGAGGAATTAATCTTGAAGTCTGGGCAACAGAGGACATATGAGCAGC	2460
Db	2401	CTGTCTTGAAGGAGGAATTAATCTTGAAGTCTGGGCAACAGAGGACATATGAGCAGC	2460
QY	2461	CAAGAATGCCCGTCTCTGTTCAAGTTAAATCAAGGATTCACCTTCTTCCCTACCAA	2520
Db	2461	CAAGAATGCCCGTCTCTGTTCAAGTTAAATCAAGGATTCACCTTCTTCCCTACCAA	2520
QY	2521	GGGAGTACCCCTCAGACCAAGGCCCAACAAAGGATTCCAAAGATTTGTTAAGGACTTAA	2580
Db	2521	GGGAGTACCCCTCAGACCAAGGCCCAACAAAGGATTCCAAAGATTTGTTAAGGACTTAA	2580
QY	2581	AAGCCCAAGGCTTAGTAAACCATGCACTAATCCCTGCGAGTAAATCCGTTAGTGATGAG	2640
Db	2581	AAGCCCAAGGCTTAGTAAACCATGCACTAATCCCTGCGAGTAAATCCGTTAGTGATGAG	2640
QY	2641	GAGGACACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGGATCTCAGGATTAATCAATG	2700
Db	2641	GAGGACACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGGATCTCAGGATTAATCAATG	2700
QY	2701	AGGCGGTGTCTCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTTCCCAATAC	2760
Db	2701	AGGCGGTGTCTCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTTCCCAATAC	2760
QY	2761	CAGAGGACAGAGTGGTTTACATCTCTGGAACCTTMAAGATGCTTCTCTGATCCCTG	2820
Db	2761	CAGAGGACAGAGTGGTTTACATCTCTGGAACCTTMAAGATGCTTCTCTGATCCCTG	2820
QY	2821	TACATCTGACTCTCAATTTCTTGTGCTTTGAGAGATACCTTCAAAACCCATCTCAAC	2880
Db	2821	TACATCTGACTCTCAATTTCTTGTGCTTTGAGAGATACCTTCAAAACCCATCTCAAC	2880
QY	2881	TCACCTGGACTTTTATACCCCAAGGTTTCAGGATAGYCCCATCTATTTGGCCAGGCAT	2940
Db	2881	TCACCTGGACTTTTATACCCCAAGGTTTCAGGATAGYCCCATCTATTTGGCCAGGCAT	2940
QY	2941	TAGCCCAAGACTTTGAGYCARITYMTATACCTGGAACCTTCTGCTCTGTAAGTGGATG	3000
Db	2941	TAGCCCAAGACTTTGAGYCARITYMTATACCTGGAACCTTCTGCTCTGTAAGTGGATG	3000
QY	3001	ATTTACTTTTTCGCTTTCAGAAACCTTGTGTCATCAAGCCCAAGGCTCTTTMA	3060
Db	3001	ATTTACTTTTTCGCTTTCAGAAACCTTGTGTCATCAAGCCCAAGGCTCTTTMA	3060
QY	3061	ATTTCTCTGCTGCTGCTACAGGTTTCCAAACARARCTCTCTCTCTCACAGC	3120
Db	3061	ATTTCTCTGCTGCTGCTACAGGTTTCCAAACARARCTCTCTCTCTCACAGC	3120
QY	3121	AGGTTAAATCTAGGCTTAARATTTATCAAAAGGACCAAGGCCCTCAGTGGAGAAATY	3180
Db	3121	AGGTTAAATCTAGGCTTAARATTTATCAAAAGGACCAAGGCCCTCAGTGGAGAAATY	3180
QY	3181	TCCAGCTTACTGCTTCTCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTTCTG	3240
Db	3181	TCCAGCTTACTGCTTCTCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTTCTG	3240

Qy	3241	GCR	TAA	YAG	YTTCTG	CCGA	AAW	TG	AT	TCC	CAG	GTW	TGG	CR	AA	TAG	CC	AG	GY	CA	T	T	A		3300	
Db	3241	GCR	TAA	YAG	YTTCTG	CCGA	AAW	TG	AT	TCC	CAG	GTW	TGG	CR	AA	TAG	CC	AG	GY	CA	T	T	A		3300	
Qy	3301	WAT	AC	AST	AA	T	TAAG	AAA	CT	CAGA	AG	CCA	TAC	CC	CA	T	T	T	ART	TA	GA	T	GA	Y	3360	
Db	3301	WAT	AC	AST	AA	T	TAAG	AAA	CT	CAGA	AG	CCA	TAC	CC	CA	T	T	T	ART	TA	GA	T	GA	Y	3360	
Qy	3361	GYM	RA	AG	TGG	CT	T	T	T	T	C	CAG	G	CC	C	T	T	T	T	T	T	T	T	T	3420	
Db	3361	GYM	RA	AG	TGG	CT	T	T	T	T	C	CAG	G	CC	C	T	T	T	T	T	T	T	T	T	3420	
Qy	3421	TG	CA	AC	CR	G	G	G	CA	AG	CT	T	T	T	T	T	T	T	T	T	T	T	T	T	3480	
Db	3421	TG	CA	AC	CR	G	G	G	CA	AG	CT	T	T	T	T	T	T	T	T	T	T	T	T	T	3480	
Qy	3481	GT	C	T	T	T	AC	AG	R	T	CC	BA	GG	AY	GA	G	T	T	T	T	T	T	T	T	3540	
Db	3481	GT	C	T	T	AC	AG	R	T	CC	BA	GG	AY	GA	G	T	T	T	T	T	T	T	T	T	3540	
Qy	3541	GA	T	AG	T	GG	CA	AG	GG	T	T	GC	Y	CA	T	T	T	T	T	T	T	T	T	T	3600	
Db	3541	GA	T	AG	T	GG	CA	AG	GG	T	T	GC	Y	CA	T	T	T	T	T	T	T	T	T	T	3600	
Qy	3601	GT	A	T	C	T	GA	AG	C	AG	T	TA	AA	TA	TA	CA	GG	R	AG	AT	C	T	T	A	3660	
Db	3601	GT	A	T	C	GA	AG	C	AG	T	TA	AA	TA	TA	TA	CA	GG	R	AG	AT	C	T	T	A	3660	
Qy	3661	GT	GA	Y	RG	CA	T	AC	T	GC	T	AA	AG	GA	CA	T	T	G	CT	GT	C	A	CA	C	3720	
Db	3661	GT	GA	Y	RG	CA	T	AC	T	GC	T	AA	AG	GA	CA	T	T	G	CT	GT	C	A	CA	C	3720	
Qy	3721	TR	T	C	AG	CT	CT	AT	T	AC	T	T	GA	RG	SC	AG	T	GC	T	GR	AC	T	T	G	3780	
Db	3721	TR	T	C	AG	CT	CT	AT	T	AC	T	T	GA	RG	SC	AG	T	GC	T	GR	AC	T	T	G	3780	
Qy	3781	CC	AG	Y	C	N	C	A	T	T	C	T	CC	AG	CA	AT	GA	GA	AA	GA	T	A	A	A	3840	
Db	3781	CC	AG	Y	C	N	C	A	T	T	C	T	CC	AG	CA	AT	GA	GA	AA	GA	T	A	A	A	3840	
Qy	3841	T	C	T	CA	A	A	C	C	T	A	T	AG	CC	A	C	T	CG	AG	G	G	A	C	T	3900	
Db	3841	T	C	T	CA	A	A	C	C	T	A	T	AG	CC	A	C	T	CG	AG	G	G	A	C	T	3900	
Qy	3901	CA	A	C	T	T	CA	T	CA	T	CG	AA	GT	CC	T	T	T	G	A	GA	AA	AG	CA	T	3960	
Db	3901	CA	A	C	T	T	CA	T	CA	T	CG	AA	GT	CC	T	T	T	G	A	GA	AA	AG	CA	T	3960	
Qy	3961	AG	T	GG	T	C	AG	T	GA	T	AA	T	Y	T	GA	AA	GT	AA	T	CC	CT	CA	C	T	4020	
Db	3961	AG	T	GG	T	C	AG	T	GA	T	AA	T	Y	T	GA	AA	GT	AA	T	CC	CT	CA	C	T	4020	
Qy	4021	G	C	T	R	G	C	A	A	C	T	A	T	AG	CC	A	C	T	CG	AG	G	G	A	C	T	4080
Db	4021	G	C	T	R	G	C	A	A	C	T	A	T	AG	CC	A	C	T	CG	AG	G	G	A	C	T	4080
Qy	4081	TAT	AT	A	T	AC	AG	AC	T	CT	PA	T	AT	GC	TY	AC	T	AG	T	CC	AT	GC	CC	AT	4140	
Db	4081	TAT	AT	A	T	AC	AG	AC	T	CT	PA	T	AT	GC	TY	AC	T	AG	T	CC	AT	GC	CC	AT	4140	
Qy	4141	AGA	AG	G	GA	A	T	T	CC	T	TA	CT	CY	AG	R	GA	CA	C	CT	PA	CA	M	A	C	4200	
Db	4141	AGA	AG	G	GA	A	T	T	CC	T	TA	CT	CY	AG	R	GA	CA	C	CT	PA	CA	M	A	C	4200	
Qy	4201	ATT	AT	T	AT	Y	T	G	S	C	WG	T	AC	AG	AA	C	CT	AB	AG	G	T	G	G	M	4260	
Db	4201	ATT	AT	T	AT	Y	T	G	S	C	WG	T	AC	AG	AA	C	CT	AB	AG	G	T	G	G	M	4260	
Qy	4261	NA	AG	GA	AA	G	RA	AG	G	GA	A	T	A	S	A	AG	GA	Y	T	GC	CA	AG	CA	K	4320	
Db	4261	NA	AG	GA	AA	G	RA	AG	G	GA	A	T	A	S	A	AG	GA	Y	T	GC	CA	AG	CA	K	4320	
Qy	4321	TG	CA	AG	G	C	AG	CA	CC	CT	CA	T	T	AG	AA	A	T	GC	T	T	A	T	T	A	4380	

Db	4321	TGCAAGCGCAGGACCCCTTCACATTAGAAATGCTTATTAAACTTCCTCTTAGTATAGGGTAATCC	4380
Qy	4381	CTTTCCGGGAAAACAAGCCCCAGTACTACGACGAGAGAAAAGAAATGGGGAACTCTCACGAGG	4440
Db	4381	CTTCCGGGAAAACAAGCCCCAGTACTACGACGAGAGAAAAGAAATGGGGAACTCTCACGAGG	4440
Qy	4441	CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAAAATACTTTTGCTGGAAC	4500
Db	4441	CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAAGGGAAAAAATACTTTTGCTGCAAC	4500
Qy	4501	TATCCAATGGAAATTACTTTAAACCCTTCATCAAACCTTTCACTTAGGCATCATAGGCAC	4560
Db	4501	TATCCAATGGAAATTACTTTAAACCCTTCATCAAACCTTTCACTTAGGCATCATAGGCAC	4560
Qy	4561	CCATCARATGGCCAAATCATTTATTACTGCACAGGCCCTTTTCAAACCTATCAAGCARAT	4620
Db	4561	CCATCARATGGCCAAATCATTTATTACTGCACAGGCCCTTTTCAAACCTATCAAGCARAT	4620
Qy	4621	AKTCCAGGGCTGTGAARKTGTGCCARARAATAATCCCCTGCCTYATCGCAAGCTCCTTC	4680
Db	4621	AKTCCAGGGCTGTGAARKTGTGCCARARAATAATCCCCTGCCTYATCGCAAGCTCCTTC	4680
Qy	4681	AGGARAAACAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCACAAG	4740
Db	4681	AGGARAAACAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCACAAG	4740
Qy	4741	CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGTARATACTTTTCAGGGTGGCA	4800
Db	4741	CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGTARATACTTTTCAGGGTGGCA	4800
Qy	4801	RAGGCCTTCCTCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCATCTAGTTTCATGAA	4860
Db	4801	RAGGCCTTCCTCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCATCTAGTTTCATGAA	4860
Qy	4861	ATAATTTCCAGATTCGGACTTTCCCAGGGCTTTACAGAGTGA CAATAGCCCTGCTTTCCAG	4920
Db	4861	ATAATTTCCAGATTCGGACTTTCCCAGGGCTTTACAGAGTGA CAATAGCCCTGCTTTCCAG	4920
Qy	4921	GCCACAGTAAACCAGGGAGTATCCCAGGGCTTTAGGTATACGATATCACTTACACTCGGCC	4980
Db	4921	GCCACAGTAAACCAGGGAGTATCCCAGGGCTTTAGGTATACGATATCACTTACACTCGGCC	4980
Qy	4981	TGAAGGCCACAGTCCTCAGGAAAGTTCGAGAAATGAATGAAAYATCTAAAGGACATCTA	5040
Db	4981	TGAAGGCCACAGTCCTCAGGAAAGTTCGAGAAATGAATGAAAYATCTAAAGGACATCTA	5040
Qy	5041	AAAAAGCAAACCCAGGAAAACCCACCTCA CATGGCCCTGYTCTGTGTGCTATAGCCTTAAAA	5100
Db	5041	AAAAAGCAAACCCAGGAAAACCCACCTCA CATGGCCCTGYTCTGTGTGCTATAGCCTTAAAA	5100
Qy	5101	AGAACTCGAACCTTTCCC AAAAGCAGNCTTAGGCCCATACGAAATGCTGTATGGAAG	5160
Db	5101	AGAACTCGAACCTTTCCC AAAAGCAGNCTTAGGCCCATACGAAATGCTGTATGGAAG	5160
Qy	5161	CCCTTCATACCAATGACCTTGCTGTGACCCCAAGACAGCAACTTAGTTGTGCAGACATCA	5220
Db	5161	CCCTTCATACCAATGACCTTGCTGTGACCCCAAGACAGCAACTTAGTTGTGCAGACATCA	5220
Qy	5221	CCTCCTTTAGCCAAATATCAACAAGTTCTTTAAACATTTACAAGGAACCTATCCCTGAGAAG	5280
Db	5221	CCTCCTTTAGCCAAATATCAACAAGTTCTTTAAACATTTACAAGGAACCTATCCCTGAGAAG	5280
Qy	5281	AGGGAAGAAGCACTATTTCCACCWWGTGACATGTTATTAGTCAAGTCCCTTCYTCTAATT	5340
Db	5281	AGGGAAGAAGCACTATTTCCACCWWGTGACATGTTATTAGTCAAGTCCCTTCYTCTAATT	5340
Qy	5341	CCCCATCCCTTAGATACATCTCTGGGAAGGACCCCTACCCAGCTATTTTATYTACCCCAACTG	5400
Db	5341	CCCCATCCCTTAGATACATCTCTGGGAAGGACCCCTACCCAGCTATTTTATYTACCCCAACTG	5400
Qy	5401	CGGTTAAGTGGCTGGAGTGAGTCTTTGGATACATCACTTCAGTCAAAATCCTCGGATAC	5460

Db 5401 CGGTTAAAGTGCGTAGTGAGTCTTTGGATACATCACACTTGAGTCAAATCCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACCTGAAATCCAGAGACAAACGCTAGCTATTCTGTGAAACCTCTAGAGG 5520  
Db 5461 TGCCAAAGGAACCTGAAATCCAGAGACAAACGCTAGCTATTCTGTGAAACCTCTAGAGG 5520  
Qy 5521 ATTTGGCGCTGCTCTTCAAACAAACAAACAGGAGGAAAGTAACTAAATCATAAATCCCCC 5580  
Db 5521 ATTTGGCGCTGCTCTTCAAACAAACAAACAGGAGGAAAGTAACTAAATCATAAATCCCCC 5580  
Qy 5581 ATGSGCCTCCCTTATCATATATTTTCTCTKTAAGTCTTTTACCTCTTTTCACTCTCACT 5640  
Db 5581 ATGSGCCTCCCTTATCATATATTTTCTCTKTAAGTCTTTTACCTCTTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCGCGTATGACACGAGTACCTCCCTTACCMAGAGTTTCTATGAGGA 5700  
Db 5641 GCACCCCTCCATGCGCGTATGACACGAGTACCTCCCTTACCMAGAGTTTCTATGAGGA 5700  
Qy 5701 ATGCAGCGTCCCGGAAATATTGATGCCCACTCGTATAGGAGTCTTTSTAAGGGAACCCC 5760  
Db 5701 ATGCAGCGTCCCGGAAATATTGATGCCCACTCGTATAGGAGTCTTTSTAAGGGAACCCC 5760  
Qy 5761 ACCTTCACTGCCCCACACCATATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATG 5820  
Db 5761 ACCTTCACTGCCCCACACCATATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATG 5820  
Qy 5821 CATGCCAAATCACTATTTATGGAAGGAAATATTAATCCTAGTTGCTTGGAGGACTT 5880  
Db 5821 CATGCCAAATCACTATTTATGGAAGGAAATATTAATCCTAGTTGCTTGGAGGACTT 5880  
Qy 5881 GGAGTCACTGCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGAGAAAACATGTTAAAGAAAGTAACTCTCCCAACTCACCGSGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGAGAAAACATGTTAAAGAAAGTAACTCTCCCAACTCACCGSGGTACAT 6000  
Qy 6001 GGCACCTCTAGCCCCCTCAAAGGACTAGTCTCTCAAACCTACATGAAACCCCTCCGTAAC 6060  
Db 6001 GGCACCTCTAGCCCCCTCAAAGGACTAGTCTCTCAAACCTACATGAAACCCCTCCGTAAC 6060  
Qy 6061 CATACTCGCTGTGAAGCCTATTTAATACACCCCTCACTGGCTCCATGAGGTCTCGGCC 6120  
Db 6061 CATACTCGCTGTGAAGCCTATTTAATACACCCCTCACTGGCTCCATGAGGTCTCGGCC 6120  
Qy 6121 CAAAACCCCTACTAATCTGTGATGATGCTCCCTCGAACTTTCAAGCCATATGTTCAATC 6180  
Db 6121 CAAAACCCCTACTAATCTGTGATGATGCTCCCTCGAACTTTCAAGCCATATGTTCAATC 6180  
Qy 6181 CCTGTACCTGAACAAATGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTATGTA 6240  
Db 6181 CCTGTACCTGAACAAATGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTATGTA 6240  
Qy 6241 GGACCTCTGTTCCTAATSTGGAATTAACCCATACCTCAAACCTCACTGTGTAAATTT 6300  
Db 6241 GGACCTCTGTTCCTAATSTGGAATTAACCCATACCTCAAACCTCACTGTGTAAATTT 6300  
Qy 6301 AGCAATACTACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCCAACAA 6360  
Db 6301 AGCAATACTACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCCAACAA 6360  
Qy 6361 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGATACCTCAGGCTATCGTTGTTTG 6420  
Db 6361 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGATACCTCAGGCTATCGTTGTTTG 6420  
Qy 6421 AATGGCTCTTCAAGATCTATGCTTCTCTCAATTTAGTGCCCCCYATGRCACATCTAC 6480  
Db 6421 AATGGCTCTTCAAGATCTATGCTTCTCTCAATTTAGTGCCCCCYATGRCACATCTAC 6480  
Qy 6481 ACTGAAACAGATTTATACAGTTATGTCATATCTAAGCCCCCGCAACAAAGAGTACCCATT 6540  
Db 6481 ACTGAAACAGATTTATACAGTTATGTCATATCTAAGCCCCCGCAACAAAGAGTACCCATT 6540

Qy 6541 CTTCTCTTTTCTTATAGGAGCAGAGTCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATC 6600  
Db 6541 CTTCTCTTTTCTTATAGGAGCAGAGTCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATC 6600  
Qy 6601 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
Db 6601 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
Qy 6661 GTGCCGACTCCCTGGTGCACCTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT 6720  
Db 6661 GTGCCGACTCCCTGGTGCACCTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT 6720  
Qy 6721 CRAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAAACCTGTTATTTTAA 6780  
Db 6721 CRAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAAACCTGTTATTTTAA 6780  
Qy 6781 GGGGGAAGATGCTGTTATTTATGTTAATCAATCGGAATCGTCACTGAGAAGTTTAAAGAA 6840  
Db 6781 GGGGGAAGATGCTGTTATTTATGTTAATCAATCGGAATCGTCACTGAGAAGTTTAAAGAA 6840  
Qy 6841 ATTCSAGATCGAATACAAACGTAKAGCAGAGAGCTTCTGAAACACTGGACCTCTGGGGCTC 6900  
Db 6841 ATTCSAGATCGAATACAAACGTAKAGCAGAGAGCTTCTGAAACACTGGACCTCTGGGGCTC 6900  
Qy 6901 CTGAGCCRAATGAGTCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Db 6901 CTGAGCCRAATGAGTCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Qy 6961 CTACTCCTCTTTGGACCTCTGATCTTTTACCTCTTGTAACTTTGTCTCTCCAGAAATC 7020  
Db 6961 CTACTCCTCTTTGGACCTCTGATCTTTTACCTCTTGTAACTTTGTCTCTCCAGAAATC 7020  
Qy 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATCAGTCCAAGACTAGATCTACCCGAGA 7080  
Db 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATCAGTCCAAGACTAGATCTACCCGAGA 7080  
Qy 7081 CCCTGGACCGGCTGTGTAGCCACGATCTGTATGTTTAAATGACATCAAAGGACACCCCTCT 7140  
Db 7081 CCCTGGACCGGCTGTGTAGCCACGATCTGTATGTTTAAATGACATCAAAGGACACCCCTCT 7140  
Qy 7141 GAGGAATCTCAGCTGCACAACTCTATACGCCCCCAATTCAGAGAGAGCAGTTAGAGC 7200  
Db 7141 GAGGAATCTCAGCTGCACAACTCTATACGCCCCCAATTCAGAGAGAGCAGTTAGAGC 7200  
Qy 7201 GGTSGTCGCGCAACCTCCCAACAGCACTTAGGTCTTCTGTTGAGATGGGCACTGAGA 7260  
Db 7201 GGTSGTCGCGCAACCTCCCAACAGCACTTAGGTCTTCTGTTGAGATGGGCACTGAGA 7260  
Qy 7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCVTAAGCCTTAGSTGGGAAGGTG 7320  
Db 7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCVTAAGCCTTAGSTGGGAAGGTG 7320  
Qy 7321 ACCAATCCACCTTTTAAACACGCGGCTTGCACTTAGYTCAACCTTGACCAATCAGAGAG 7380  
Db 7321 ACCAATCCACCTTTTAAACACGCGGCTTGCACTTAGYTCAACCTTGACCAATCAGAGAG 7380  
Qy 7381 CTCACCTAAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTC 7440  
Db 7381 CTCACCTAAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTC 7440  
Qy 7441 MTGAGAGCAGCAGCAGGAGGGAACAATGATCGGATATAAACCCCAAGTCTTCGAGCCCGCAA 7500  
Db 7441 MTGAGAGCAGCAGCAGGAGGGAACAATGATCGGATATAAACCCCAAGTCTTCGAGCCCGCAA 7500  
Qy 7501 CGGCAACCCCTTTGGGTCCCTTCTGTTGATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Db 7501 CGGCAACCCCTTTGGGTCCCTTCTGTTGATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Qy 7561 TCTATTAAATCTTGCARCTGCR 7582  
Db 7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 2  
 US-10-717-580-11  
 ; Sequence 11, Application US/10717580  
 ; Publication No. US20040176314A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BESEME, Frederic  
 ; APPLICANT: BLOND, Jean-Luc  
 ; APPLICANT: BOUTON, Olivier  
 ; APPLICANT: MANDRAND, Bernard  
 ; APPLICANT: MALLET, Francois  
 ; APPLICANT: PERRON, Hervé  
 ; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES  
 ; TITLE OF INVENTION: WITH PREGNANCY DISORDERS  
 ; FILE REFERENCE: 105045  
 ; CURRENT APPLICATION NUMBER: US/10/717,580  
 ; CURRENT FILING DATE: 2003-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/446,024A  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/FR98/01442  
 ; PRIOR FILING DATE: 1998-07-06  
 ; PRIOR APPLICATION NUMBER: FR 97/08815  
 ; PRIOR FILING DATE: 1997-07-07  
 ; NUMBER OF SEQ ID NOS: 53  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 11  
 ; LENGTH: 7582  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (198)..(198)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (307)..(307)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1309)..(1309)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1331)..(1331)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (2213)..(2213)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (2398)..(2398)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (3787)..(3787)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4115)..(4115)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4261)..(4261)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; US-10-717-580-11  
 Query Match 100.0%; Score 7582; DB 9; Length 7582;  
 Best Local Similarity 100.0%; Pred. No. 0;

	Matches	7582;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	CAACAATCGGATATATAAACCCAGGCAATTCGAGCTGGGCAACAGCAGCCCTTTGGGTCC	60							
DB	1	CAACAATCGGATATATAAACCCAGGCAATTCGAGCTGGGCAACAGCAGCCCTTTGGGTCC	60							
QY	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTATTCCTCTATTAAATCTTGCACACTGCA	120							
DB	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTATTCCTCTATTAAATCTTGCACACTGCA	120							
QY	121	CTCTTCTGGTCCATGTTTCTTAACGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180							
DB	121	CTCTTCTGGTCCATGTTTCTTAACGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180							
QY	181	TGTTTGGCCACCAACCCGACCTGCGGCTGACTCCCATCCCTCTGATCTCTGAGGCTGC	240							
DB	181	TGTTTGGCCACCAACCCGACCTGCGGCTGACTCCCATCCCTCTGATCTCTGAGGCTGC	240							
QY	241	CGCTGTGCTCCTGATCCAGCGGCGCCATTTGCCGCTCCCAATTTGGGCTAAAGGCTTGC	300							
DB	241	CGCTGTGCTCCTGATCCAGCGGCGCCATTTGCCGCTCCCAATTTGGGCTAAAGGCTTGC	300							
QY	301	CATTGTNCTGCAACGCTAAGTGCCTGGGTTGTTCTAATTGAGCTGAACACTANTCACT	360							
DB	301	CATTGTNCTGCAACGCTAAGTGCCTGGGTTGTTCTAATTGAGCTGAACACTANTCACT	360							
QY	361	GGGTTCCATGTTCTCTTCTGTGACCCAGGCTTCTTAATAAATCTTAACACTTACCACA	420							
DB	361	GGGTTCCATGTTCTCTTCTGTGACCCAGGCTTCTTAATAAATCTTAACACTTACCACA	420							
QY	421	TGGCCCAAGATTCCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTTCAGAGATAC	480							
DB	421	TGGCCCAAGATTCCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTTCAGAGATAC	480							
QY	481	GARGCTTGCCACCATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTATC	540							
DB	481	GARGCTTGCCACCATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTATC	540							
QY	541	TTGGGAGCTCTGTGAGCAAGACCCCGGCTGACATTTTGGGACCAACMSRACGGACATCC	600							
DB	541	TTGGGAGCTCTGTGAGCAAGACCCCGGCTGACATTTTGGGACCAACMSRACGGACATCC	600							
QY	601	MAAGTATGGGAAACGTTCCCGCAAGACAAAAACGCGCTTAAGACGCTATTCTGGARAT	660							
DB	601	MAAGTATGGGAAACGTTCCCGCAAGACAAAAACGCGCTTAAGACGCTATTCTGGARAT	660							
QY	661	TGGGAMCAATTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTGCAGTGCC	720							
DB	661	TGGGAMCAATTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTGCAGTGCC	720							
QY	721	GCCTGGCACTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACTCTTTTG	780							
DB	721	GCCTGGCACTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACTCTTTTG	780							
QY	781	TAGAAAAGGCAATTCGAGTGAAGTCCCATAGTACAACTTTCTTTTCAATTAAGACAAA	840							
DB	781	TAGAAAAGGCAATTCGAGTGAAGTCCCATAGTACAACTTTCTTTTCAATTAAGACAAA	840							
QY	841	CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCC	900							
DB	841	CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCC	900							
QY	901	CTATCCGACATCCCGACTCCTTCCCAATTAAGGACCCCTTCAACCCCAATGG	960							
DB	901	CTATCCGACATCCCGACTCCTTCCCAATTAAGGACCCCTTCAACCCCAATGG	960							
QY	961	TCCAAAAGGAGATAGACAAAAAGGGTAAACAGTGAACCAAGAGTGCCTATTTCCCAAT	1020							
DB	961	TCCAAAAGGAGATAGACAAAAAGGGTAAACAGTGAACCAAGAGTGCCTATTTCCCAAT	1020							
QY	1021	TATGACCCCTCCCAAGCAGTGGGAGAGAGATTTCCGCCACGACAGTGTGCTT	1080							
DB	1021	TATGACCCCTCCCAAGCAGTGGGAGAGAGATTTCCGCCACGACAGTGTGCTT	1080							









QY	5461	TGCCAAAGGAACCTGAAATCCAGGACAAACGCTAGCTATTCTGTGAACCTCTAGAGG	5520
Db	5461		
QY	5521	ATTTGGCCCTGCTCTTCAAAACAACACAGGAGGAAAGTAACATAAAATCATAAATCCCCC	5580
Db	5521		
QY	5581	ATGSCCTCCCTATCATATATTTCTCTKTAAGTTSTTTTACCTSTTTCACTCTCACT	5640
Db	5581		
QY	5641	GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTTACCMAGAGTTTCTATGGAGA	5700
Db	5641		
QY	5701	ATGCAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGAGTCTTTSTAAGGGAACCCCC	5760
Db	5701		
QY	5761	ACCTTCACTGCCACACCCATATATGCCCCATCGTATAGAGTCTTTSTAAGGGAACCCCC	5820
Db	5761		
QY	5821	CATGCAAACTACTCATTTATTTGGACAGGAAATGATTAATCCTAGTTGTCTGAGGACTT	5880
Db	5821		
QY	5881	GGAGTCACTGCTGTGTGGACTTACTTCCACCAACTGGTATGTCTGATGGGGTGAGTT	5940
Db	5881		
QY	5941	CAAGATCAGCAAGAGAGAAAACATGTAAAGAAAGTAATCTCCAACTCACCSGGGTACAT	6000
Db	5941		
QY	6001	GGCACTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTAATGAAGACCTCCGTACC	6060
Db	6001		
QY	6061	CATACTCGCTGTAGGCTTATTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGC	6120
Db	6061		
QY	6121	CAAAACCTCTACTAATCTGTGGATATGCTCCCTCGAACTTCARGCCATATGTTCAATC	6180
Db	6121		
QY	6181	CTGTGTACCTGAACAATGGAAACAATTCAGCACAGAAATAAACACCACTTCGGTTTTAGTA	6240
Db	6181		
QY	6241	GGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT	6300
Db	6241		
QY	6301	AGCAATACTACATACACAACCACTCCCAATGATCAGGTGGGTAACTCTCTCCACACAA	6360
Db	6301		
QY	6361	ATAGTCTGCTACCCCTCAGGAATTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6420
Db	6361		
QY	6421	AATGGCTCTCAAGATCTATGTGCTCTCTATCTTATGTGCCCCCATGRCATCTAC	6480
Db	6421		
QY	6481	ACTGAACAAGATTTATACAGTTATGTATCATATCTAAGCCCGGCAACAAAGAGTACCATT	6540
Db	6481		

QY	6541	CTTCTCTTTTGTATACGAGCAGGAGTGTAGGTGCACTAGGTACTGGCAATTGCGGTATC	6600
Db	6541		
QY	6601	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	6660
Db	6601		
QY	6661	GTGCGCGACTCCCTGTGTCACTTGTGCAAGATCAACTTAACCTCTAGCAGCAGTAGTCTT	6720
Db	6661		
QY	6721	CRAAATCGAGAGCTTTAGACTYCTAAACGCTGAPAGAGGGGGAACCTGTTATTTTAA	6780
Db	6721		
QY	6781	GGGGAAGATGCTGTTATTATTTATTAATCAATCCGGAATCGTCACTCAGAAAGTTAAGAA	6840
Db	6781		
QY	6841	ATTCSAGATCGAATACAACGTAKAGCAGARGAGCTTCGAAACAACCTGGGCGCTC	6900
Db	6841		
QY	6901	CTCAGCCRAATGAGTGCCTGGATTCTCCCTCTTAGGACCTCTAGCAGCTATAATTTG	6960
Db	6901		
QY	6961	CTACTCTCTTTGGACCTGTATCTTTTACCTCTTTTAACTTTCTCTCCAGATC	7020
Db	6961		
QY	7021	GAAGCTGTAAACTACAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCCGAGA	7080
Db	7021		
QY	7081	CCCTGGACCGGCTGTGTAGCCCAAGATGCAGTCCAAGACTAAGATCTACCCGAGA	7140
Db	7081		
QY	7141	GAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAGCAGTTAGAC	7200
Db	7141		
QY	7201	GCTGTGCGCCCACTCCCAACAGCACTTAGGTTTCTCTGTAGATGGGACCTGAGA	7260
Db	7201		
QY	7261	GACGAGCTAGCTGGATTTCCTAGGCTGAYTAAGAAATCCYTAAGCCTAGSTGGGAAGTG	7320
Db	7261		
QY	7321	ACCAATCCACCTTTAAACACCGGGCTTGCACCTTAGYTCACACCTGACCAATCAGAG	7380
Db	7321		
QY	7381	CTCACTAAATGCTAATTAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTC	7440
Db	7381		
QY	7441	MTGAGAGCACAGCAGGAGGACAATGATCGGGATATAACCCCAAGTTCGAGCCCGCAA	7500
Db	7441		
QY	7501	CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTCAATGCTATTTTAC	7560
Db	7501		
QY	7561	TCCTATTAAATCTTGCARCTGCR	7582
Db	7561		

```

RESULT 3
US-10-450-763-4430
; Sequence 4430, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CI3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 4430
; LENGTH: 7466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1378)..(2271)
; OTHER INFORMATION: 84% homologous to multiple sclerosis associated retrovirus
; OTHER INFORMATION: polyprotein,accession number AF009668,Smith-Waterman Score=1279.
US-10-450-763-4430

```

Query Match	12.9%;	Score 979;	DB 10;	Length 7466;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1269;	Conservative 0;	Mismatches 1;	Indels 2;	Gaps 2;
Qy	1476	GAATTACTCAATGATGTCCACCATAAACACAGGGGAAGGAAGAAATCCCTACTGCGCTTTC	1535	
Db	5407	GAATTACTCAATGATGTCCACCATAAACACAGGGGAAGGAAGAAATCCCTACTGCGCTTTC	5466	
Qy	1536	TGGAGAGACTAAGGGAGGCGATTGAGGAAGCGTGCTCTCTGTCACTCGACTCTTCTTGAAG	1595	
Db	5467	TGGAGAGACTAAGGGAGGCGATTGAGGAAGCGTGCTCTCTGTCACTCGACTCTTCTTGAAG	5526	
Qy	1596	GCCAACTAATCTTAAAGGTTAAGTTTATCACTCAGTCAGCTGCAGACATTTAGAAAAA	1655	
Db	5527	GCCAACTAATCTTAAAGCGTTAAGTTTATCACTCAGTCAGCTGCAGACATTTAGAAAAA	5586	
Qy	1656	TCAAAGTCTGCCGTAGGCCCGGAGCAAACTTAGAAAACCTATTGAAACCTTTGGCAAC	1715	
Db	5587	TCAAAGTCTGCCGTAGGCCCGGAGCAAACTTAGAAAACCTATTGAACTTTGGCAAC	5646	
Qy	1716	GGTTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAAACAGGACAAACGGGATTTAAAA	1775	
Db	5647	GGTTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAAACAGGACAAACGGGATTTAAAA	5706	
Qy	1776	AGGCCACGGCTTTAGTCATGACCCCTCAGGCNAGTGGAATTTGGAGGCTCTGGAAAA	1835	
Db	5707	AGGCCACGGCTTTAGTCATGACCCCTCAGGCNAGTGGAATTTGGAGGCTCTGGAAAA	5766	
Qy	1836	AAAGCTGGGCAAAATTGAATGCCTTAATAGGGCTTCCTTCCTCCAGTGGCGGTCTACA	1895	
Db	5767	AAAGCTGGGCAAAATTGAATGCCTTAATAGGGCTTCCTTCCTCCAGTGGCGGTCTACA	5826	
Qy	1896	TTAAAAAGATTTGTCGAAGTAGAAGTGAAGCGCGCCCTTTGTCGATGCGCCCTTATT	1955	
Db	5827	TTAAAAAGATTTGTCGAAGTAGAAGTGAAGCGCGCCCTTTGTCGATGCGCCCTTATT	5886	
Qy	1956	GGAATCACTGGAAGGCCACATGCCCCAGGGGACAAAGGTCTTTTGTAGTCAGAGGCCA	2015	
Db	5887	GGAATCACTGGAAGGCCACATGCCCCAGGGGACAAAGGTCTTTTGTAGTCAGAGGCCA	5946	
Qy	2016	ACCAAGATGATCCAGCAGCAGACTGAGGGTGCTTGGGGCAAGGCCCATCCCATGCCAT	2075	
Db	5947	ACCAAGATGATCCAGCAGCAGACTGAGGGTGCTTGGGGCAAGGCCCATCCCATGCCAT	6006	

## RESULT 4

```

US-10-632-793-29
; Sequence 29, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 2372
; TYPE: DNA

```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1191)..(1191)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1213)..(1213)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2089)..(2089)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2274)..(2274)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-29

Query Match      11.7%; Score 888; DB 8; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGGAGGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGCG 1566
DB 1383 GGGAGGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGCG 1442

QY 1567 TGCCTCTCTGTCACTGACTTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1626
DB 1443 TGCCTCTCTGTCACTGACTTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1502

QY 1627 TCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGCGCCGGAGCAAAAC 1686
DB 1503 TCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGCGCCGGAGCAAAAC 1562

QY 1687 TTAGAAACCTATTGAACCTTGCAACVTCGGTTTTTTTATAATAGACATCAGGAGGAGCAG 1746
DB 1563 TTAGAAACCTATTGAACCTTGCAACVTCGGTTTTTTTATAATAGACATCAGGAGGAGCAG 1622

QY 1747 GCGGAACAGGACAAACCGGATTAATAAAAGGCGCCAGCTTGTAGTCATGACCCCTCAGGCA 1806
DB 1623 GCGGAACAGGACAAACCGGATTAATAAAAGGCGCCAGCTTGTAGTCATGACCCCTCAGGCA 1682

QY 1807 AGTGACTTTGGAGGCTCTGGAAGAGGAAAGCTGGGCAAAATGTAATGCTTAATAGGC 1866
DB 1683 AGTGACTTTGGAGGCTCTGGAAGAGGAAAGCTGGGCAAAATGTAATGCTTAATAGGC 1742

QY 1867 TTGCTTCCAGTCGGTCTCAAGGACACTTAAAAAGATTTGTCAGAGTAGAGTAGCC 1926
DB 1743 TTGCTTCCAGTCGGTCTCAAGGACACTTAAAAAGATTTGTCAGAGTAGAGTAGCC 1802

QY 1927 GCCCTTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1986
DB 1803 GCCCTTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1862

QY 1987 ACAAGGCTTTTGTAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGTG 2046
DB 1863 ACAAGGCTTTTGTAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGTG 1922

QY 2047 CCTGGGGCAGGCCCATCCATGCCATCACCCCTCAGAGGCCCTGGGTATGCTTGACCAT 2106
DB 1923 CCTGGGGCAGGCCCATCCATGCCATCACCCCTCAGAGGCCCTGGGTATGCTTGACCAT 1982

QY 2107 TGAGGGCCAGGAAGGTGTCTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACTCTTCG 2166
DB 1983 TGAGGGCCAGGAAGGTGTCTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACTCTTCG 2042

QY 2167 TCCCGGACAACTGTCTCTCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGACGGGCAGT 2226
DB 2043 TCCCGGACAACTGTCTCTCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGACGGGCAGT 2102

QY 2227 CACTAGATATCTTCTCCAGCACTAAGTTATGAACCTGGGAGCTTTATCTTTTTCACAT 2286
DB 1383 GGGAGGGGAAGAAATCCTACTGCTTCTTGAGAGACTAAGGAGGCATTGAGGAAGCG 1566
DB 1383 GGGAGGGGAAGAAATCCTACTGCTTCTTGAGAGACTAAGGAGGCATTGAGGAAGCG 1442
```

```
DB 2103 CACTAGATATCTTCTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTTTCACAT 2162
QY 2287 GCTTTTCTAATTATGCTTGAAGCCCACTACCTTCTGTAGGGAGAGACATTCTTAGCAAAA 2346
DB 2163 GCCTTTCTAATTATGCTTGAAGCCCACTACCTTCTGTAGGGAGAGACATTCTTAGCAAAA 2222
QY 2347 GCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTGTGTCCTGCT 2406
DB 2223 GCAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTGTGTCCTGCT 2282
QY 2407 TGAGGAAGGAATTAATCTGGAAGTCTGGGCAACAGGAAGGACATATATGGACGCCAAAGA 2466
DB 2283 TGAGGAAGGAATTAATCTGGAAGTCTGGGCAACAGGAAGGACATATATGGACGCCAAAGA 2342
QY 2467 ATGCCCGTCTCTGTTCAAGTTAAACTAAAGG 2496
DB 2343 ATGCCCGTCTCTGTTCAAGTTAAACTAAAGG 2372
```

## RESULT 5

```
US-10-717-580-10
; Sequence 10, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DIS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1191)..(1191)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1213)..(1213)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2089)..(2089)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2274)..(2274)
; OTHER INFORMATION: n = any nucleotide
US-10-717-580-10
```

Query Match 11.7%; Score 888; DB 9; Length 2372;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1567 TGCCTCTCTGTCCACCTGACTCTCTGAAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC 1626  
Db 1443 TGCCTCTCTGTCCACCTGACTCTCTGAGGCCACTAATCTTAAAGCGTAAGTTTATCAC 1502  
QY 1627 TCAGTCAGCTGCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTAGGCGCCGAGCAAAAC 1686  
Db 1503 TCAGTCAGCTGCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTAGGCGCCGAGCAAAAC 1562  
QY 1687 TTAGAAACCCATTGAATCTTGCACTGCGCAATCTGCGTTTATTAATAGAGATCAGGAGAGCAG 1746  
Db 1563 TTAGAAACCCATTGAATCTTGCACTGCGCAATCTGCGTTTATTAATAGAGATCAGGAGAGCAG 1622  
QY 1747 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1806  
Db 1623 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1682  
QY 1807 AGTGACCTTTTGAGGCTCTGGAAGAGGAAAGCTGGGCAAAATTTGAATGCTTAATAGGCG 1866  
Db 1683 AGTGACCTTTTGAGGCTCTGGAAGAGGAAAGCTGGGCAAAATTTGAATGCTTAATAGGCG 1742  
QY 1867 TTGCTTCCAGTCGGCTCTACAAGGACACTTTTAAAAAGATTGTCCAAAGTAGAAGTAAGCC 1926  
Db 1743 TTGCTTCCAGTCGGCTCTACAAGGACACTTTTAAAAAGATTGTCCAAAGTAGAAGTAAGCC 1802  
QY 1927 GCCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1986  
Db 1803 GCCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1862  
QY 1987 ACAAGAGTCTTTTGAGTCAGAAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGGTG 2046  
Db 1863 ACAAGAGTCTTTTGAGTCAGAAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGGTG 1922  
QY 2047 CTTGGGCGAAGCCCATCCATGCACTCAAGGCTTGGGTATGCTTCAACCAT 2106  
Db 1923 CTTGGGCGAAGCCCATCCATGCACTCAAGGCTTGGGTATGCTTCAACCAT 1982  
QY 2107 TGAGGCGCAGGAAGTGTCTCTCGACACTGGTGGGTCTCTTAGTCTTACTCTTCTG 2166  
Db 1983 TGAGGCGCAGGAAGTGTCTCTCGACACTGGTGGGTCTCTTAGTCTTACTCTTCTG 2042  
QY 2167 TCCCGGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGTCNTAAGACGGGCAGT 2226  
Db 2043 TCCCGGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGTCNTAAGACGGGCAGT 2102  
QY 2227 CACTAGATATCTTTTCCAGCCACTAAAGTTATGAATCTGGGAGCTTTTATCTTTTCAAT 2286  
Db 2103 CACTAGATATCTTTTCCAGCCACTAAAGTTATGAATCTGGGAGCTTTTATCTTTTCAAT 2162  
QY 2287 GCTTTTCTAAATATGCTTGAAAGCCCACTACTCTGTTAGGAGAGACATTTCTAGCAAA 2346  
Db 2163 GCTTTTCTAAATATGCTTGAAAGCCCACTACTCTGTTAGGAGAGACATTTCTAGCAAA 2222  
QY 2347 GCAGGGGCCATTTATACCTGTAACATAGGAGAGGAACACCCGTTGTTGTCNCCCTGCT 2406  
Db 2223 GCAGGGGCCATTTATACCTGTAACATAGGAGAGGAACACCCGTTGTTGTCNCCCTGCT 2282  
QY 2407 TGAGGAAGGAATTAATCTTGAAGTCTGGCCAAACAGAGGACAAATATGACGAGCCAAAGA 2466  
Db 2283 TGAGGAAGGAATTAATCTTGAAGTCTGGCCAAACAGAGGACAAATATGACGAGCCAAAGA 2342  
QY 2467 ATGCCCGTCTGTTCAAGTTAACTAAAGG 2496  
Db 2343 ATGCCCGTCTGTTCAAGTTAACTAAAGG 2372

RESULT 6  
US-11-028-539-1  
; Sequence 1, Application US/11028539  
; Publication No. US20050118573A1  
; GENERAL INFORMATION:  
; APPLICANT: ALLIEL, Patrick  
; APPLICANT: PERIN, Jean-Pierre  
; APPLICANT: RIEGER, Francois

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU  
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS  
; FILE REFERENCE: 200936USOPCT  
; CURRENT APPLICATION NUMBER: US/11/028,539  
; CURRENT FILING DATE: 2005-01-05  
; PRIOR APPLICATION NUMBER: US/09/719,554  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
; PRIOR FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 2599  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-028-539-1

Query Match 10.3%; Score 779; DB 13; Length 2599;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
QY 4802 AGGCTTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCATAGTTTCATGAAA 4861  
Db 151 AGGCTTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCATAGTTTCATGAAA 210  
QY 4862 TAATTTCCAGATTCCGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTGCTTCCAGG 4921  
Db 211 TAATTTCCAGATTCCGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTGCTTCCAGG 270  
QY 4922 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCT 4981  
Db 271 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCT 330  
QY 4982 GAAGGCCACAGTCTCTAGGGAAGGTTCGAGAAAAATGAATGAAYACTCAAAAGACATCTAA 5041  
Db 331 GAAGGCCACAGTCTCTAGGGAAGGTTCGAGAAAAATGAATGAAYACTCAAAAGACATCTAA 390  
QY 5042 AAAAGCAAAACCCAGGAACCCACCTCATCGGCTGTCTGTTCGCTATAGCTTAAAAA 5101  
Db 391 AAAAGCAAAACCCAGGAACCCACCTCATCGGCTGTCTGTTCGCTATAGCTTAAAAA 450  
QY 5102 GAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATAGCAATGCTGTATGGAAGGC 5161  
Db 451 GAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATAGCAATGCTGTATGGAAGGC 510  
QY 5162 CTTTCATAACCAATGACCTTGTGTGACCCAAAGACAGCCAACTTAGTTGCGAGACATCAC 5221  
Db 511 CTTTCATAACCAATGACCTTGTGTGACCCAAAGACAGCCAACTTAGTTGCGAGACATCAC 570  
QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTACAAGGAACCTATCCCTGAGAGA 5281  
Db 571 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTACAAGGAACCTATCCCTGAGAGA 630  
QY 5282 GGGAAAAAAGCTTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 5341  
Db 631 GGGAAAAAAGCTTATTTCCACCTTGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 690  
QY 5342 CCATTCCTTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTTATYACCCCACTGC 5401  
Db 691 CCATTCCTTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTTATCTACCCCACTGC 750  
QY 5402 GGTTHAAGTGGCTGGAGTGGAGTCTTGGATATACATCACATTGAGTCAAACTCTGGATACT 5461  
Db 751 GGTTHAAGTGGCTGGAGTGGAGTCTTGGATATACATCACATTGAGTCAAACTCTGGATACT 810  
QY 5462 GCCAAAGGAACCTGAAAAATCCAGGAGACAAACGCTAGTATTCTCTGTGAACCTCTAGAGA 5521  
Db 811 GCCAAAGGAACCTGAAAAATCCAGGAGACAAACGCTAGTATTCTCTGTGAACCTCTAGAGA 870  
QY 5522 TTTGCGCTGCTCTTCAAAACAAACCAAGGAGGAAAGTAACATAAATCATAAATCCCCCA 5581  
Db 871 TTTGCGCTGCTCTTCAAAACAAACCAAGGAGGAAAGTAACATAAATCATAAATCCCCCA 929

QY 5582 TGGSCCTCCCTATCATATTTTCTCTKTAATGTTSTTTTACCCCTSTTTTCACTCTCACTG 5641  
Db 930 TGGCCCTCCCTATCATATTTTCTCTTACTGTTCTTTTACCCCTCTTTTCACTCTCACTG 989  
QY 5642 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTCYACMAGAGTTTCTATGGAGAA 5701  
Db 990 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTCYACMAGAGTTTCTATGGAGAA 1049  
QY 5702 TGCAGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGACCCCA 5761  
Db 1050 TGCAGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGACCCCA 1109  
QY 5762 CTTTCACTGCCACACCCTATGTCGCCCACTGCTATCACTCTGCCACCTTTTGCATGC 5821  
Db 1110 CTTTCACTGCCACACCCTATGTCGCCCACTGCTATCACTCTGCCACCTTTTGCATGC 1169  
QY 5822 ATGCAATATCTATTTATGACAGGAAAAATGATTAATCTCTAGTGTGCTGGAGGACTG 5881  
Db 1170 ATGCAATATCTATTTATGACAGGAAAAATGATTAATCTCTAGTGTGCTGGAGGACTG 1229  
QY 5882 GAGTCACTGCTCTGGACTTCTTACCCAACTGCTATGCTGTATGCTGGGGTGGAGTTC 5941  
Db 1230 GAGTCACTGCTCTGGACTTCTTACCCAACTGCTATGCTGTATGCTGGGGTGGAGTTC 1289  
QY 5942 AAGATCAGGCAAGAAAAATGTAAGAAAGTAAATCTCCCACTCACCGSGGTACATG 6001  
Db 1290 AAGATCAGGCAAGAAAAATGTAAGAAAGTAAATCTCCCACTCACCGSGGTACATG 1349  
QY 6002 GCACCTCTAGCCCTTACAAGGACTAGATCTCTCAAACTTACATGAACCTCCGTACCC 6061  
Db 1350 GCACCTCTAGCCCTTACAAGGACTAGATCTCTCAAACTTACATGAACCTCCGTACCC 1409  
QY 6062 ATACTCGCTGTGAAGCTATTTAATACACCTCTCTGCGGTCCATGAGGTCTCGGCC 6121  
Db 1410 ATACTCGCTGTGAAGCTATTTAATACACCTCTCTGCGGTCCATGAGGTCTCGGCC 1469  
QY 6122 AAAACCTTACTACTGTTGATATGCTCCCTGAACTTCAGGCCATATGTTTCAATCC 6181  
Db 1470 AAAACCTTACTACTGTTGATATGCTCCCTGAACTTCAGGCCATATGTTTCAATCC 1529  
QY 6182 CTGTACTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTAG 6241  
Db 1530 CTGTACTGAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTAG 1589  
QY 6242 GACCTCTGTTTCCAAATSTGGAATAAACCATACCTCAAACTCAGCTGTGTAAATTTA 6301  
Db 1590 GACCTCTGTTTCCAAATSTGGAATAAACCATACCTCAAACTCAGCTGTGTAAATTTA 1649  
QY 6302 GCAATACTACATACAAACCACTCCCAATGCATCAGGTGGGTAACTCTCCACACAA 6361  
Db 1650 GCAATACTACATACAAACCACTCCCAATGCATCAGGTGGGTAACTCTCCACACAA 1709  
QY 6362 TAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTGA 6421  
Db 1710 TAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTGTACCTCAGCCTATCGTTGTTGA 1769  
QY 6422 ATGGCTCTTCAGAACTATGTGCTTCTCTCAATTTAGTGCCCTCCATGRCATCTACA 6481  
Db 1770 ATGGCTCTTCAGAACTATGTGCTTCTCTCAATTTAGTGCCCTCCATGRCATCTACA 1829  
QY 6482 CTGAACAAGATTTATACAGTTATGTCAATCTTAAGCCCGCAACAAAGAGTACCAATTC 6541  
Db 1830 CTGAACAAGATTTATACAGTTATGTCAATCTTAAGCCCGCAACAAAGAGTACCAATTC 1889  
QY 6542 TTCCTTTTGTATAGGAGCAGAGTCTAGGTGCACTAGGTACTGCGATTTGGGGTATCA 6601  
Db 1890 TTCCTTTTGTATAGGAGCAGAGTCTAGGTGCACTAGGTACTGCGATTTGGGGTATCA 1949  
QY 6602 CAACTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGACGGG 6661  
Db 1950 CAACTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGACGGG 2009  
QY 6662 TCCCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC 6721

Db 2010 TGCCCGACTCCCTGGTCACTTGAAGATCACTTAATCCCTAGCAGCAGTAGTCTCCTC 2069  
RESULT 7  
US-09-854-867-21  
; Sequence 21, Application US/09854867  
; Publication No. US20030224356A1  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL H  
; APPLICANT: ROGAN, PETER K  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATION  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/09/854,867  
; CURRENT FILING DATE: 2003-05-08  
; NUMBER OF SEQ ID NOS: 613  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 8523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat region  
; LOCATION: (1)..(8523)  
; OTHER INFORMATION: herv17  
US-09-854-867-21  
Query Match 10.3%; Score 779; DB 3; Length 8523;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
QY 4802 AGGCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCATGAAA 4861  
Db 6071 AGGCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCATGAAA 6130  
QY 4862 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTACATAGCCCTGCTTTCAGG 4921  
Db 6131 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTACATAGCCCTGCTTTCAGG 6190  
QY 4922 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATCTGCGCCT 4981  
Db 6191 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATCTGCGCCT 6250  
QY 4982 GAAGGCCACAGTCTCAGGGAAGGTCGAGAAAAATGAATGAAAYACTCAAGGACATCTAA 5041  
Db 6251 GAAGGCCACAGTCTCAGGGAAGGTCGAGAAAAATGAATGAAAYACTCAAGGACATCTAA 6310  
QY 5042 AAAAGCAACCCAGGAACCCACCTCACATGGCGTCTGTTGCTTATAGCCTTAAAAA 5101  
Db 6311 AAAAGCAACCCAGGAACCCACCTCACATGGCGTCTGTTGCTTATAGCCTTAAAAA 6370  
QY 5102 GAATCTGCAACTTCCCAAAAGCAGGACTTACCCCATACGAAATGCTGTATGGAAGGC 5161  
Db 6371 GAATCTGCAACTTCCCAAAAGCAGGACTTACCCCATACGAAATGCTGTATGGAAGGC 6430  
QY 5162 CTTTCATAACCAATGACCTTGTGTTGACCCCAAGCAGCACTTAGTTTGCAGACATCAC 5221  
Db 6431 CTTTCATAACCAATGACCTTGTGTTGACCCCAAGCAGCACTTAGTTTGCAGACATCAC 6490  
QY 5222 CTCCTTAGCCAAATATCAAGTTCTTAAAAATTAATCAAGGAACTATCCCTGAGAAGA 5281  
Db 6491 CTCCTTAGCCAAATATCAAGTTCTTAAAAATTAATCAAGGAACTATCCCTGAGAAGA 6550  
QY 5282 GGGAAAGAACTATTCCACCCWGTGACATGATGATAGTCAAGTCCCTTCCTTAATTC 5341  
Db 6551 GGGAAAGAACTATTTCACCCCTTGTGACATGATATAGTCAAGTCCCTTCCTTAATTC 6610  
QY 5342 CCCATCCCTAGATACATCTGGGAAGGACCTTACCCAGTCAATTTATATACCCCACTGC 5401  
Db 6611 CCCATCCCTAGATACATCTGGGAAGGACCTTACCCAGTCAATTTATCTACCCCACTGC 6670  
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAATCTCGGATCT 5461

Db 6671 GGTAAAGTGGCTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATACT 6730  
Qy 5462 GCCAAGGAACTGAAATCCAGGAGACACGCTAGCTATTCCTGTGAACCTCTAGAGGA 5521  
Db 6731 GCCAAGGAACTGAAATCCAGGAGACACGCTAGCTATTCCTGTGAACCTCTAGAGGA 6790  
Qy 5522 TTGTGGCCCTGCTCTTCAACAACAACACGAGAGAAAGTAACATAAATCATAAATCCCCA 5581  
Db 6791 TTGTGGCCCTGCTCTTCAACAACAACACGAGAGAAAGTAACATAAATCATAAATCCCCA 6849  
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTPASTGTTSTTTAYACCCSTTTTCACTCTCACTG 5641  
Db 6850 TGGCCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCTCTTCACTCTCACTG 6909  
Qy 5642 GACCCCTCCATGCGCTGTATGACGAGTAGCTCCCTCYACCMAGATTTCTATGGAGAA 5701  
Db 6910 CACCCCTCCATGCGCTGTATGACGAGTAGCTCCCTTACCAAGAGTTTCTATGGAGAA 6969  
Qy 5702 TGCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 5761  
Db 6970 TCGAGGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCA 7029  
Qy 5762 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTCATGC 5821  
Db 7030 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTCATGC 7089  
Qy 5822 ATGCAAACTACTATTTATGGACAGGAAATAATGATTAATCTAGTTGTCTCGAGGACTTG 5881  
Db 7090 ATGCAAACTACTATTTATGGACAGGAAATAATGATTAATCTAGTTGTCTCGAGGACTTG 7149  
Qy 5882 GAGTCACCTGCTGTGGACTTCTCACCCAACTGGTATGCTCATGGGGTGGAGTTC 5941  
Db 7150 GAGTCACCTGCTGTGGACTTCTCACCCAACTGGTATGCTCATGGGGTGGAGTTC 7209  
Qy 5942 AAGATCAGGCAAGAGAAACATGTAAGAAAGATTAATCTCCCAACTCACCSGGGTACATG 6001  
Db 7210 AAGATCAGGCAAGAGAAACATGTAAGAAAGATTAATCTCCCAACTCACCSGGGTACATG 7269  
Qy 6002 GCACCTTAGCCCTTACAAGAGCTAGATCTCTCAAACTAATGAAGACCTTCGGTACCC 6061  
Db 7270 GCACCTTAGCCCTTACAAGAGCTAGATCTCTCAAACTAATGAAGACCTTCGGTACCC 7329  
Qy 6062 ATACTCGCTGCTAGGCTATTTAATACCACTCTGCGGCTCCATGAGGTCTCGGCC 6121  
Db 7330 ATACTCGCTGCTAGGCTATTTAATACCACTCTGCGGCTCCATGAGGTCTCGGCC 7389  
Qy 6122 AAAACCTACTAACTGTGTGATATGCTCCCTGAACTTTCARGGCATATGTTTCAATCC 6181  
Db 7390 AAAACCTACTAACTGTGTGATATGCTCCCTGAACTTTCARGGCATATGTTTCAATCC 7449  
Qy 6182 CTGTACTTGAAACAATGGAAACAATTCAGACAGAAATAAACAACCACTTCCTCGTTTGTAG 6241  
Db 7450 CTGTACTTGAAACAATGGAAACAATTCAGACAGAAATAAACAACCACTTCCTCGTTTGTAG 7509  
Qy 6242 GACCTCTGTTTCCAAATGGAATAACCATACCTCAAACTCCACCTGCTGTAATAATTTA 6301  
Db 7510 GACCTCTGTTTCCAAATGGAATAACCATACCTCAAACTCCACCTGCTGTAATAATTTA 7569  
Qy 6302 GCAATACTACATACACAACAACCACTCCCAATGATCAGGTGGGTAACTCTCCACACAAA 6361  
Db 7570 GCAATACTACATACACAACAACCACTCCCAATGATCAGGTGGGTAACTCTCCACACAAA 7629  
Qy 6362 TAGTCTGCTTACCTCAGGAATAATTTTGTCTGTGGTACCTCAGGCTATCGTTTGTGA 6421  
Db 7630 TAGTCTGCTTACCTCAGGAATAATTTTGTCTGTGGTACCTCAGGCTATCGTTTGTGA 7689  
Qy 6422 ATGGCTCTTCAGAACTATGCTCTCTCATCTTGTAGTCCCCVATGRCATCTACA 6481  
Db 7690 ATGGCTCTTCAGAACTATGCTCTCTCATCTTGTAGTCCCCCTATGACCATCTACA 7749  
Qy 6482 CTGAACAAGATTTATACAGTTATGTATCTAAGCCCGGCAACAAAAGAGTACCCATTC 6541  
Db 7750 CTGAACAAGATTTATACAGTTATGTATCTAAGCCCGGCAACAAAAGAGTACCCATTC 7809

Qy 6542 TTCCTTTGTTATTAGGACGAGGTGCTAGTGCACCTAGGTACTGGCATTTGGCGGTATCA 6601  
Db 7810 TTCCTTTGTTATTAGGACGAGGTGCTAGTGCACCTAGGTACTGGCATTTGGCGGTATCA 7869  
Qy 6602 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGG 6661  
Db 7870 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGG 7929  
Qy 6662 TCGCGACTCCCTGCTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCTTC 6721  
Db 7930 TCGCGACTCCCTGCTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCCTTC 7989

RESULT 8

US-10-786-970A-21  
; Sequence 21, Application US/10786970A  
; Publication No. US2005006449A1

GENERAL INFORMATION:

; APPLICANT: JOAN, KNOLL  
; APPLICANT: ROGAN, PETER  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/10/786,970A  
; CURRENT FILING DATE: 2004-02-24  
; PRIOR APPLICATION NUMBER: US/09/573,080  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 479  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 21

; LENGTH: 8523

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: repeat region

; LOCATION: (1)..(8523)

; OTHER INFORMATION: herv17

; PUBLICATION INFORMATION:

; PUBLICATION INFORMATION:

; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A

; TITLE: Prototypic sequences for human repetitive DNA

; JOURNAL: Journal of Molecular Evolution

; VOLUME: 35

; ISSUE: 4

; PAGES: 286-291

; DATE: 1992-10-

; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)

; DATABASE ENTRY DATE:

; DATABASE ENTRY DATE: 1996-01-26

US-10-786-970A-21

Query Match 10.3%; Score 779; DB 10; Length 8523;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

Qy 4802 AGGCCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAGGCATAGTTCATGAAA 4861  
Db 6071 AGGCCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAGGCATAGTTCATGAAA 6130  
Qy 4862 TAAATCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 4921  
Db 6131 TAAATCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 6190  
Qy 4922 CCACAGTAACCCAGGAGTATCCCAAGGCTTAGGTATACGATATCACTTACACTGCGCCT 4981  
Db 6191 CCACAGTAACCCAGGAGTATCCCAAGGCTTAGGTATACGATATCACTTACACTGCGCCT 6250  
Qy 4982 GAAGGCCACAGTCTTCAGGAGGTCGAGAAATCAATCAAAVACTCAAGACATCTAA 5041  
Db 6251 GAAGGCCACAGTCTTCAGGAGGTCGAGAAATCAATCAAAVACTCAAGACATCTAA 6310  
Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGCTGTGCTATAGCCTTAAAAA 5101

Db 6311 AAAAGCAAAACCAGGAAACCCACCTCACATGGCCCTGCTCTGTTCCTATAGCCTTAAAAA 6370  
 Qy 5102 GAATCTGCAATTTTCCCAAAAAGCAGGACTTACCCCATACGAAATCTGTATGGAAGGC 5161  
 Db 6371 GAATCTGCAATTTTCCCAAAAAGCAGGACTTACCCCATACGAAATCTGTATGGAAGGC 6430  
 Qy 5162 CTTTCATAACCAATGACCTTGTGCTTACCCAGAGAGAGCACTTATGTCGACAGCATCAC 5221  
 Db 6431 CTTTCATAACCAATGACCTTGTGCTTACCCAGAGAGAGCACTTATGTCGACAGCATCAC 6490  
 Qy 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTAACAAGAACCTTATCCCTGAGAGA 5281  
 Db 6491 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTAACAAGAACCTTATCCCTGAGAGA 6550  
 Qy 5282 GGGAAAGAACTATTCACCCWGTGACATGATGATTAAGTCAAGTCCCTTCYCTCAATTC 5341  
 Db 6551 GGGAAAGAACTATTCACCCWGTGACATGATGATTAAGTCAAGTCCCTTCYCTCAATTC 6610  
 Qy 5342 CCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCAATTTATATACCCCACTGC 5401  
 Db 6611 CCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCAATTTATATACCCCACTGC 6670  
 Qy 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAATCAAAATCTTGGATACT 5461  
 Db 6671 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAATCAAAATCTTGGATACT 6730  
 Qy 5462 GCCAAAGAACTGAAATCCAGAGACAAACGCTAGCTATTTCTGTGAACTCTAGAGGA 5521  
 Db 6731 GCCAAAGAACTGAAATCCAGAGACAAACGCTAGCTATTTCTGTGAACTCTAGAGGA 6790  
 Qy 5522 TTTGGCGCTCTCTTCAAAACAAACAGAGAGAAAGTAACTAAATCATAAATCCCCCA 5581  
 Db 6791 TTTGGCGCTCTCTTCAAAACAAACAGAGAGAAAGTAACTAAATCATAAATCCCCCA 6849  
 Qy 5582 TGGSCCTCCCTATCATATTTTCTCTKTASTGTSTTTTACCCCTSTTTTCACTCTCACTG 5641  
 Db 6850 TGGSCCTCCCTATCATATTTTCTCTKTASTGTSTTTTACCCCTSTTTTCACTCTCACTG 6909  
 Qy 5642 CACCCCTCCATCGCGTGTATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGGAGAA 5701  
 Db 6910 CACCCCTCCATCGCGTGTATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGGAGAA 6969  
 Qy 5702 TGCAGCGTCCGGAATATTTGATGCGCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 5761  
 Db 6970 TGCAGCGTCCGGAATATTTGATGCGCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 7029  
 Qy 5762 CTTTCACTGCCACACCCATATGCGCGCACTGCTATCACTTGCACCTTTTGCATGC 5821  
 Db 7030 CTTTCACTGCCACACCCATATGCGCGCACTGCTATCACTTGCACCTTTTGCATGC 7089  
 Qy 5822 ATGCAAACTACTTATTTGGACAGGAAATGATTAATCTAGTGTCTTGGAGGACTTG 5881  
 Db 7090 ATGCAAACTACTTATTTGGACAGGAAATGATTAATCTAGTGTCTTGGAGGACTTG 7149  
 Qy 5882 GAGTCACTGTCTGTGGACTTACTTCAACCAACTGATGCTGTATGCTGATGGGGTGGAGTTC 5941  
 Db 7150 GAGTCACTGTCTGTGGACTTACTTCAACCAACTGATGCTGTATGCTGATGGGGTGGAGTTC 7209  
 Qy 5942 AAGATCAGGCAAGAGAAACATGTAAGAAGTATCTCCCAACTCACCGGGGTACATG 6001  
 Db 7210 AAGATCAGGCAAGAGAAACATGTAAGAAGTATCTCCCAACTCACCGGGGTACATG 7269  
 Qy 6002 GCACCTTAGCCCTTACAAAGGACTAGTCTCTCAAAACTTACATGAACCCCTCCGTACCC 6061  
 Db 7270 GCACCTTAGCCCTTACAAAGGACTAGTCTCTCAAAACTTACATGAACCCCTCCGTACCC 7329  
 Qy 6062 ATACTCGCTGTGAAGCTATTTTAAATACCACTCTGAGGCTCCATGAGGCTCTCGGCC 6121  
 Db 7330 ATACTCGCTGTGAAGCTATTTTAAATACCACTCTGAGGCTCCATGAGGCTCTCGGCC 7389  
 Qy 6122 AAAACCTACTAATCTGTGGATATGCTCCCTGAACTTTTCAAGCCATATGTTTCAATCC 6181  
 Db 7390 AAAACCTACTAATCTGTGGATATGCTCCCTGAACTTTTCAAGCCATATGTTTCAATCC 7449

Qy 6182 CTGTACTCTGAACAATGGAACAACCTTACGACACAGAAATAAACACCACTTCCGTTTTAGTAG 6241  
 Db 7450 CTGTACTCTGAACAATGGAACAACCTTACGACACAGAAATAAACACCACTTCCGTTTTAGTAG 7509  
 Qy 6242 GACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTTCAAACTCCTGCTGTGTAAATTTA 6301  
 Db 7510 GACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTTCAAACTCCTGCTGTGTAAATTTA 7569  
 Qy 6302 GCAATACTACTACATACACAACCACTCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAA 6361  
 Db 7570 GCAATACTACTACATACACAACCACTCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAA 7629  
 Qy 6362 TAGTCTGCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCTGTTGTTGA 6421  
 Db 7630 TAGTCTGCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCTGTTGTTGA 7689  
 Qy 6422 ATGGCTCTTCAGAACTATGTGCTCTCTCATCTTCTAGTGGCCCTTATGRCCTATCTACA 6481  
 Db 7690 ATGGCTCTTCAGAACTATGTGCTCTCTCATCTTCTAGTGGCCCTTATGRCCTATCTACA 7749  
 Qy 6482 CTGAACAAGATTTATACAGTTATGTATCATATCTAAGCCCGCAACAAAGAGTACCCATTC 6541  
 Db 7750 CTGAACAAGATTTATACAGTTATGTATCATATCTAAGCCCGCAACAAAGAGTACCCATTC 7809  
 Qy 6542 TTCCTTTTGTATPAGGAGCAGGAGTCTAGTGCATCTAGTACTTGGCGGTATCA 6601  
 Db 7810 TTCCTTTTGTATPAGGAGCAGGAGTCTAGTGCATCTAGTACTTGGCGGTATCA 7869  
 Qy 6602 CAACTCTACTCTAGTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
 Db 7870 CAACTCTACTCTAGTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 7929  
 Qy 6662 TGGCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGAGTACTCTTC 6721  
 Db 7930 TGGCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGAGTACTCTTC 7989

RESULT 9

US-11-028-539-3  
 ; Sequence 3, Application US/11028539  
 ; Publication No. US20050118573A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALLIEL, Patrick  
 ; APPLICANT: PERIN, Jean-Pierre  
 ; APPLICANT: RIEGER, Francois  
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HT  
 ; FILE REFERENCE: 200936US0PCT  
 ; CURRENT APPLICATION NUMBER: US/11/028,539  
 ; PRIOR FILING DATE: 2005-01-05  
 ; PRIOR APPLICATION NUMBER: US/09/719,554  
 ; PRIOR FILING DATE: 2001-01-18  
 ; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 10499  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-11-028-539-3

Query Match 10.3%; Score 779; DB 13; Length 10499;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 Qy 4802 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTATTAAGGCACTAGTTCATGAAA 4861  
 Db 7101 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTATTAAGGCACTAGTTCATGAAA 7160  
 Qy 4862 TAAATCCCAAGATCGGACTTCCCGAGGCTTACAGAGTCACAATAGCCCTGCTTCCAGG 4921



Db 7161 TAATCCAGATTCCGAGTTCCTCCGAGGCTTACAGAGTGCATAATAGCCCTTCCTCCAGG 7220  
 Qy 4922 CCACAGTAAACCCAGGAGATCCAGGCGTTAGGTATACGATATCACTTACACCTGGGCT 4981  
 Db 7221 CACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACCTGGGCT 7280  
 Qy 4982 GAAGGCCACAGTCTCAGGGAAGGTGAGAAAATGAATGAAAYACTCAAGAGACATCTAA 5041  
 Db 7281 GAAGGCCACAGTCTCAGGGAAGGTGAGAAAATGAATGAAACACTCAAGAGACATCTAA 7340  
 Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGTGCTATAGCTTAAAAA 5101  
 Db 7341 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGTGCTATAGCTTAAAAA 7400  
 Qy 5102 GAATCTGCAACTTCCCAAAAAGCAGGACTTAGCCCATACAAATGCTGTATGAAGGC 5161  
 Db 7401 GAATCTGCAACTTCCCAAAAAGCAGGACTTAGCCCATACAAATGCTGTATGAAGGC 7460  
 Qy 5162 CTTTCAATACCAATGACCTTGTGCTTACCCCAAGACAGCAACTTAGTTGACAGATCAC 5221  
 Db 7461 CTTTCAATACCAATGACCTTGTGCTTACCCCAAGACAGCAACTTAGTTGACAGATCAC 7520  
 Qy 5222 CTCCTTAGCCAAATATCAACAAGTCTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 5281  
 Db 7521 CTCCTTAGCCAAATATCAACAAGTCTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 7580  
 Qy 5282 GGGAAAAGAACTATTCCACCCHWGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAAATTC 5341  
 Db 7581 GGGAAAAGAACTATTCCACCCTTGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAAATTC 7640  
 Qy 5342 CCATCCCTTAGATACATCTGCGGAAGGACCTACCCAGTCACTTTATATVACCCCAACTGC 5401  
 Db 7641 CCATCCCTTAGATACATCTGCGGAAGGACCTACCCAGTCACTTTATATVACCCCAACTGC 7700  
 Qy 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATACT 5461  
 Db 7701 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATACT 7760  
 Qy 5462 GCCAAAGAACTGAAAATCCAGGAGACAAACGCTAGCTATTCCTGTGAACCTCTAGAGGA 5521  
 Db 7761 GCCAAAGAACTGAAAATCCAGGAGACAAACGCTAGCTATTCCTGTGAACCTCTAGAGGA 7820  
 Qy 5522 TTTGGCGCTGCTCTTCAACACACACCCAGGAGGAAGTAACTAAATCATATAATCCCCCA 5581  
 Db 7821 TTTGGCGCTGCTCTTCAACACACACCCAGGAGGAAGTAACTAAATCATATAATCCCCCA 7879  
 Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTAAGTSTTTTACCTSTTTTCACTCTCACTG 5641  
 Db 7880 TGGSCCTCCCTTATCATATTTTCTCTTACTTCTTTTACCTCTTCACTCTCACTG 7939  
 Qy 5642 CACCCCTCCATGCGCTGTATGACAGTGTCTCCCTVACCMAGAGTTTCTATGGAGAA 5701  
 Db 7940 CACCCCTCCATGCGCTGTATGACAGTGTCTCCCTTACCAAGAGTTTCTATGGAGAA 7999  
 Qy 5702 TGCAGGCTCCGGAATATGTAGCCCATCGATAGGAGTCTTTTAAGGGAACCCCCA 5761  
 Db 8000 TGCAGGCTCCGGAATATGTAGTGGCCCATCGATAGGAGTCTTTTAAGGGAACCCCCA 8059  
 Qy 5762 CTTTCACTGCCACACCCATATGCCCCGCAACTGTCTATCACTCTGCACTCTTTGCAATGC 5821  
 Db 8060 CTTTCACTGCCACACCCATATGCCCCGCAACTGTCTATCACTCTGCACTCTTTGCAATGC 8119  
 Qy 5822 ATGCAAAATCTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCGGAGGACTTG 5881  
 Db 8120 ATGCAAAATCTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCGGAGGACTTG 8179  
 Qy 5882 GAGTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTTC 5941  
 Db 8180 GAGTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTTC 8239  
 Qy 5942 AAGATCAGGCAAGAGAAAAACATGTAAGAGAGTAACTTCCCAACTCACCSSGGGTACATG 6001  
 Db 8240 AAGATCAGGCAAGAGAAAAACATGTAAGAGAGTAACTTCCCAACTCACCSSGGGTACATG 8299

Qy 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 6061  
 Db 8300 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 8359  
 Qy 6062 ATACTCGCCTGGTAAAGCCTATTAAATACCAACCTCACCTGGGCTCCATGAGGTCTCGGCC 6121  
 Db 8360 ATACTCGCCTGGTAAAGCCTATTAAATACCAACCTCACCTGGGCTCCATGAGGTCTCGGCC 8419  
 Qy 6122 AAAACCTCTACTAATCTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC 6181  
 Db 8420 AAAACCTCTACTAATCTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC 8479  
 Qy 6182 CTGTACTCTGAACATGGAACAACTTCAGCAGAGAAATAAACCACTTCCGTTTGTAGTAG 6241  
 Db 8480 CTGTACTCTGAACATGGAACAACTTCAGCAGAGAAATAAACCACTTCCGTTTGTAGTAG 8539  
 Qy 6242 GACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACTCCTACCTGTGTAAATTTTA 6301  
 Db 8540 GACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCCTACCTGTGTAAATTTTA 8599  
 Qy 6302 GCAATCTACATACACAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 6361  
 Db 8600 GCAATCTACATACACAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 8659  
 Qy 6362 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGA 6421  
 Db 8660 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGA 8719  
 Qy 6422 ATGGCTCTTCAAGATCTATGTCTCTCAATCTTCTAGTGTCCCTCCATGACCTATCA 6481  
 Db 8720 ATGGCTCTTCAAGATCTATGTCTCTCAATCTTCTAGTGTCCCTCCATGACCTATCA 8779  
 Qy 6482 CTGAACAGATTTATACAGTTATGTCTATCTAAGCCCCGCAACAAAAGAGTACCATTTC 6541  
 Db 8780 CTGAACAGATTTATACAGTTATGTCTATCTAAGCCCCGCAACAAAAGAGTACCATTTC 8839  
 Qy 6542 TTCTTTTGTATAGGAGCAGGAGTGTAGTGTGACCTAGGTACTGGCATTTGGCGGTATCA 6601  
 Db 8840 TTCTTTTGTATAGGAGCAGGAGTGTAGTGTGACCTAGGTACTGGCATTTGGCGGTATCA 8899  
 Qy 6602 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6661  
 Db 8900 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 8959  
 Qy 6662 TCGCCGACTCCCTGGTCACTTGAAGATCACTTAACTCCCTAGCAGCAGTGTCTTC 6721  
 Db 8960 TCGCCGACTCCCTGGTCACTTGAAGATCACTTAACTCCCTAGCAGCAGTGTCTTC 9019

RESULT 10

US-09-873-367C-81  
 ; Sequence 81, Application US/09873367C  
 ; Publication No. US20030165839A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; APPLICANT: Soppet, Daniel  
 ; APPLICANT: Endress, Gregory  
 ; APPLICANT: Augustus, Meena  
 ; APPLICANT: Ebner, Reinhard  
 ; APPLICANT: Carter, Kenneth  
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
 ; FILE REFERENCE: Signature Gene Sets  
 ; CURRENT APPLICATION NUMBER: US/09/873,367C  
 ; CURRENT FILING DATE: 2003-04-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
 ; PRIOR FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/236,842  
 ; PRIOR FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/244,867  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: U.S. 60/245,084



; PRIOR FILING DATE: 2000-11-01  
 ; NUMBER OF SEQ ID NOS: 1067  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 81  
 ; LENGTH: 56093  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-873-367C-81

Query Match 10.3%; Score 779; DB 3; Length 56093;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY	4802	AGGCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCATGAAA	4861
DB	35101	AGGCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCATGAAA	35160
QY	4862	TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG	4921
DB	35161	TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG	35220
QY	4922	CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACATATCACTTACACTGCGCCT	4981
DB	35221	CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACATATCACTTACACTGCGCCT	35280
QY	4982	GAAGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAATACTCAAGAGCATCTAA	5041
DB	35281	GAAGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAATACTCAAGAGCATCTAA	35340
QY	5042	AAAAGCAAACCCAGGAAACCCACTCACATGGCTGTCTGTGTCCTTATAGCCCTTAA	5101
DB	35341	AAAAGCAAACCCAGGAAACCCACTCACATGGCTGTCTGTGTCCTTATAGCCCTTAA	35400
QY	5102	GAATCTGCAACTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGC	5161
DB	35401	GAATCTGCAACTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGC	35460
QY	5162	CCCTCATAACCAATGACCTTGTGCTGACCCAGACAGCAACTTAGTTCGACACATCAC	5221
DB	35461	CCCTCATAACCAATGACCTTGTGCTGACCCAGACAGCAACTTAGTTCGACACATCAC	35520
QY	5222	CTCCTTAGCCAAATATCAACAGTCTTAAACATTTACAGGAACCTATCCCTCGAAGA	5281
DB	35521	CTCCTTAGCCAAATATCAACAGTCTTAAACATTTACAGGAACCTATCCCTCGAAGA	35580
QY	5282	GGGAAAGAACTATTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	5341
DB	35581	GGGAAAGAACTATTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	35640
QY	5342	CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCAATTTATYTAACCCCACTGC	5401
DB	35641	CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCAATTTATYTAACCCCACTGC	35700
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTTAGTCAATCTCGGATACT	5461
DB	35701	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTTAGTCAATCTCGGATACT	35760
QY	5462	GCCAAAGGAACCTGAAATCCAGAGACACGCTAGCTATTCTCTGTGAACCTCTAGAGGA	5521
DB	35761	GCCAAAGGAACCTGAAATCCAGAGACACGCTAGCTATTCTCTGTGAACCTCTAGAGGA	35820
QY	5522	TTTTCGCGCTCTCTTCAAAACAACCCAGGAGAAAGTAACATAAAATCAATAATCCCCCA	5581
DB	35821	TTTTCGCGCTCTCTTCAAAACAACCCAGGAGAAAGTAACATAAAATCAATAATCCCCCA	35879
QY	5582	TGSGCCCTCCCTTATCATATTTTCTCTKTASTGTSTTIVACCTSTTTTCACTCTCACTG	5641
DB	35880	TGSGCCCTCCCTTATCATATTTTCTCTKTASTGTSTTIVACCTSTTTTCACTCTCACTG	35939
QY	5642	CACCCCTCCCTGCTGTATGACAGTACCTGCTTCCCTTACCAAGAGTCTTCTATGAGAA	5701
DB	35940	CACCCCTCCCTGCTGTATGACAGTACCTGCTTCCCTTACCAAGAGTCTTCTATGAGAA	35999

RESULT 11

US-10-843-641A-81

; Sequence 81, Application US/10843641A

; Publication No. US200500644541

QY	5702	TGCAGCGTCCCGGAAATATTTGATGCCCATCTGATAGGAGTCTTTSTAAGGGAACCCCA	5761
DB	36000	TGCAGCGTCCCGGAAATATTTGATGCCCATCTGATAGGAGTCTTTSTAAGGGAACCCCA	36059
QY	5762	CCTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTTGATGC	5821
DB	36060	CCTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTTGATGC	36119
QY	5822	ATGCAAAATCTCATTTATGGACAGGAAATGATTAAATCCTAGTTCTCTGGAGGACTTG	5881
DB	36120	ATGCAAAATCTCATTTATGGACAGGAAATGATTAAATCCTAGTTCTCTGGAGGACTTG	36179
QY	5882	GAGTCACTCTCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC	5941
DB	36180	GAGTCACTCTCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC	36239
QY	5942	AAGATCAGCAAGAGAGAAACATGTAAAGAAAGTAATCTCCCAACTCACCGGGTACATG	6001
DB	36240	AAGATCAGCAAGAGAGAAACATGTAAAGAAAGTAATCTCCCAACTCACCGGGTACATG	36299
QY	6002	GCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTTACATGAACCCCTCCGTACCC	6061
DB	36300	GCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTTACATGAACCCCTCCGTACCC	36359
QY	6062	ATACTCGCCTGGTAAAGCTTATTAATACACCTCACTGGGCTCAATGAGGCTCTGGCCC	6121
DB	36360	ATACTCGCCTGGTAAAGCTTATTAATACACCTCACTGGGCTCAATGAGGCTCTGGCCC	36419
QY	6122	AAAAACCTACTAACTGTTGGATATGCTCCCTGAACTTTCAGGCCATATGTTTCAATCC	6181
DB	36420	AAAAACCTACTAACTGTTGGATATGCTCCCTGAACTTTCAGGCCATATGTTTCAATCC	36479
QY	6182	CTGTACTGAAACAATGGAAACAATTTTTCAGCACAGAAATAAACACCACTTCGGTTTATG	6241
DB	36480	CTGTACTGAAACAATGGAAACAATTTTTCAGCACAGAAATAAACACCACTTCGGTTTATG	36539
QY	6242	GACCTCTGTTTCCAAATSTGGAAATAAACCCATACCTCAAACTCACCTGTGTAAAAATTA	6301
DB	36540	GACCTCTGTTTCCAAATSTGGAAATAAACCCATACCTCAAACTCACCTGTGTAAAAATTA	36599
QY	6302	GCAATACTACATACACCAACTCCCAATGATCAGTGGGTGAAGTCTCCACACAA	6361
DB	36600	GCAATACTACATACACCAACTCCCAATGATCAGTGGGTGAAGTCTCCACACAA	36659
QY	6362	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTGA	6421
DB	36660	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTGA	36719
QY	6422	ATGGCTCTTTCAGAACTATGCTTCTCTCATTTCTAGTGGCCCTATGRCATCTACA	6481
DB	36720	ATGGCTCTTTCAGAACTATGCTTCTCTCATTTCTAGTGGCCCTATGRCATCTACA	36779
QY	6482	CTGAACAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAGAGTACCCATTC	6541
DB	36780	CTGAACAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAGAGTACCCATTC	36839
QY	6542	TTCTCTTTCTTATAGGACGAGGTGCTAGGTGCACTAGTACTGTCATTTGGCGGTATCA	6601
DB	36840	TTCTCTTTCTTATAGGACGAGGTGCTAGGTGCACTAGTACTGTCATTTGGCGGTATCA	36899
QY	6602	CAACCTCTACTCAGTTCTACTACAAATATCTCAAGAACTAAATGGGGACATGGAAACGGG	6661
DB	36900	CAACCTCTACTCAGTTCTACTACAAATATCTCAAGAACTAAATGGGGACATGGAAACGGG	36959
QY	6662	TGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC	6721
DB	36960	TGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC	37019

```

; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-81

Query Match          10.3%; Score 779; DB 10; Length 56093;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCCCTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAATAAGGCACCTAGTTTCATGAAA 4861
DB 35101 AGGCCCTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAATAAGGCACCTAGTTTCATGAAA 35160

QY 4862 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 4921
DB 35161 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 35220

QY 4922 CCACAGTAACCCAGGAGTATCCCGAGGCTTACAGATATACGATATCACTTACACTCGGCCT 4981
DB 35221 CCACAGTAACCCAGGAGTATCCCGAGGCTTACAGATATACGATATCACTTACACTCGGCCT 35280

QY 4982 GAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAYACTCAAGAGGACATCTAA 5041
DB 35281 GAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAYACTCAAGAGGACATCTAA 35340

QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCAATGCGCTGTCTGTGCTATAGGCTTAAAAA 5101
DB 35341 AAAAGCAAAACCCAGGAAACCCACCTCAATGCGCTGTCTGTGCTATAGGCTTAAAAA 35400

QY 5102 GAATCTGCAACTTCCCAAAAGCAGGACTTAGGCCATACCAATGCTGTATGAAGGC 5161
DB 35401 GAATCTGCAACTTCCCAAAAGCAGGACTTAGGCCATACCAATGCTGTATGAAGGC 35460

QY 5162 CTTTCATAACCAATGACCTTGTGCTTGAACCAAGCAGCCAACTTAGTTGCGAGACATCAC 5221
DB 35461 CTTTCATAACCAATGACCTTGTGCTTGAACCAAGCAGCCAACTTAGTTGCGAGACATCAC 35520

QY 5222 CTCCTTAGCCAAATATCAACAAGTCTTAAACATTAAGGAACTTATCCCTGAGAAGA 5281
DB 35521 CTCCTTAGCCAAATATCAACAAGTCTTAAACATTAAGGAACTTATCCCTGAGAAGA 35580

QY 5282 GGGAAAAAGAACTATTCCACCOWMGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAATTC 5341

```

```

Db 35581 GGGAAAAAGAACTATTCCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCCTCTAAATTC 35640
QY 5342 CCCATCCCTAGATACATCCTCGGAAGGACCTACCGAGTCAATTTTATTTATCCCAACTGC 5401
DB 35641 CCCATCCCTAGATACATCCTCGGAAGGACCTACCGAGTCAATTTTATTTATCCCAACTGC 35700
QY 5402 GGTAAAAGTGGCTGGAGTCTTGGATACATCACAATTTAGTCAATCTTGATACT 5461
DB 35701 GGTAAAAGTGGCTGGAGTCTTGGATACATCACAATTTAGTCAATCTTGATACT 35760
QY 5462 GCCAAAGGAACTTGAAATCCAGGAGCAACGCTAGCTATTCTGTGAACCTCTAGAGGA 5521
DB 35761 GCCAAAGGAACTTGAAATCCAGGAGCAACGCTAGCTATTCTGTGAACCTCTAGAGGA 35820
QY 5522 TTTGGGCTGCTCTTCAAAACAAACGAGGAGGAAAGTAACATAAATCATAAATCCCA 5581
DB 35821 TTTGGGCTGCTCTTCAAAACAAACGAGGAGGAAAGTAACATAAATCATAAAT-CCCCA 35879
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTTCTTCTCACTCTCACTG 5641
DB 35880 TGGSCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCTCTTTCACTCTCACTG 35939
QY 5642 CACCCCTCCATGCGGCTGTATGACCAGTAGCTCCCTTACCMAGAGTTTCTTATGAGAA 5701
DB 35940 CACCCCTCCATGCGGCTGTATGACCAGTAGCTCCCTTACCMAGAGTTTCTTATGAGAA 35999
QY 5702 TGCAGGTCCTCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761
DB 36000 TGCAGGTCCTCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 36059
QY 5762 CTTTCACTGCCCCACACCATATGCCCGCACTGCTATCACTCTGCACTCTTTGCACTG 5821
DB 36060 CTTTCACTGCCCCACACCATATGCCCGCACTGCTATCACTCTGCACTCTTTGCACTG 36119
QY 5822 ATGCAATATCACTATTATTGACAGAGAAAATGATTAATCTTAGTTGTCTTGAGGACTTG 5881
DB 36120 ATGCAATATCACTATTATTGACAGAGAAAATGATTAATCTTAGTTGTCTTGAGGACTTG 36179
QY 5882 GAGTCACTGCTGTGTGGACTTATTCACCCAAACCTGATGATGCTGATGGGGTGGAGTTC 5941
DB 36180 GAGTCACTGCTGTGTGGACTTATTCACCCAAACCTGATGATGCTGATGGGGTGGAGTTC 36239
QY 5942 AAGATCAGGCAAGAGAAAACATGTAAGAGTAATCTCCCACTCACCCGGGTGACATG 6001
DB 36240 AAGATCAGGCAAGAGAAAACATGTAAGAGTAATCTCCCACTCACCCGGGTGACATG 36299
QY 6002 GCACCTCTAGCCCTTACAAAGGACTTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 6061
DB 36300 GCACCTCTAGCCCTTACAAAGGACTTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 36359
QY 6062 ATACTCGCCTGGTAAGCCTATTAAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCC 6121
DB 36360 ATACTCGCCTGGTAAGCCTATTAAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCC 36419
QY 6122 AAAACCTTACTAACTGTGGATATGCCCTCCCGTGAACCTTCARGCCATATGTTTCAATCC 6181
DB 36420 AAAACCTTACTAACTGTGGATATGCCCTCCCGTGAACCTTCARGCCATATGTTTCAATCC 36479
QY 6182 CTGTACTTGAACAATGGAACAACTTTCAGCAGAGAAATAAACAACCACTTCCGTTTGTAGTAG 6241
DB 36480 CTGTACTTGAACAATGGAACAACTTTCAGCAGAGAAATAAACAACCACTTCCGTTTGTAGTAG 36539
QY 6242 GACCTTGTGTTTCCAATSTGAAATAAACCCATACCTCAAACTCACCTGTGTGTAATAATTA 6301
DB 36540 GACCTTGTGTTTCCAATSTGAAATAAACCCATACCTCAAACTCACCTGTGTGTAATAATTA 36599
QY 6302 GCAATCTACATACACACCACTCCCAATGCATCAGGTGGTAACTCCTCCACACAAA 6361
DB 36600 GCAATCTACATACACACCACTCCCAATGCATCAGGTGGTAACTCCTCCACACAAA 36659
QY 6362 TAGTCTGCTACCTCAGGAATATTTTGTGCTGTGGTACCTTCAGCCCTATCGTTGTTGA 6421

```



Query Match	9.6%	Score 731;	DB 8;	Length 1617;
Best local Similarity	99.4%;	Pred. No. 0;		
Matches 1031;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
Qy	5685	AGAGTTTCTTATGGAGAAATGCAGCGTCCCGGAAATATTGATGATGCCCATCGTATAGGAGTCT	5744	
Db	105	AGAGTTTCTTATGGAGAAATGCAGCGTCCCGGAAATATTGATGATGCCCATCGTATAGGAGTCT	164	
Qy	5745	TTSTAAGGGAAACCCCACTTCACTGCCCAACCCATATGCCCCGCAACTGCTATCACTC	5804	
Db	165	TTCTAAGGAAACCCCACTTCACTGCCCAACCCATATGCCCCGCAACTGCTATCACTC	224	
Qy	5805	TGCCACTCTTTGCATGATGCAAAATCTCAATTTATGGACAGGAAAAATGATTAATCCTAG	5864	
Db	225	TGCCACTCTTTGCATGATGCAAAATCTCAATTTATGGACAGGAAAAATGATTAATCCTAG	284	
Qy	5865	TTGTCTCTGGAGGACTTGGAGTCACTGTCGTGTGGACTTACTTCAACCAAACTGGTATGTC	5924	
Db	285	TTGTCTCTGGAGGACTTGGAGTCACTGTCGTGTGGACTTACTTCAACCAAACTGGTATGTC	344	
Qy	5925	TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	5984	
Db	345	TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	404	
Qy	5985	ACTCACGSGGTATACATGGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA	6044	
Db	405	ACTCACGSGGTATACATGGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA	464	
Qy	6045	TGAACCTCTGTTACCCATCTCGCTGGTAAGCTATTTAATACCAACCTCACTGGGCT	6104	
Db	465	TGAACCTCTGTTACCCATCTCGCTGGTAAGCTATTTAATACCAACCTCACTGGGCT	524	
Qy	6105	CCATGAGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCCCTCCCTGAACTTCAR	6164	
Db	525	CCATGAGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCCCTCCCTGAACTTCAG	584	
Qy	6165	GCCATATGTTTCAATCCCTGTACTGAACAATGGAAACAACCTTCAGCACAGAAATAAACAC	6224	
Db	585	GCCATATGTTTCAATCCCTGTACTGAACAATGGAAACAACCTTCAGCACAGAAATAAACAC	644	
Qy	6225	CAC TTCGTTTTAGTAGGACCTCTGTTTTCGAATSTGGAATAAACCATACCTCAAACTT	6284	
Db	645	CAC TTCGTTTTAGTAGGACCTCTGTTTTCGAATSTGGAATAAACCATACCTCAAACTT	704	
Qy	6285	CACCTGTATAAATTTAGCAATACATACACAACAACCTCCCAATGCAATGCACTCAGTGGGT	6344	
Db	705	CACCTGTATAAATTTAGCAATACATACACAACAACCTCCCAATGCAATGCACTCAGTGGGT	764	
Qy	6345	AACTCCTCCCAACAATAAGTCTGCCTACCTCAGGAATAATTTTTTCTGTGTGTAACCTC	6404	
Db	765	AACTCCTCCCAACAATAAGTCTGCCTACCTCAGGAATAATTTTTTCTGTGTGTAACCTC	824	
Qy	6405	AGCCTATCGTTTGAATGSCCTCTTCCAGAACTATATGCTTCTCTCATTTCTTAGTGCC	6464	
Db	825	AGCCTATCGTTTGAATGSCCTCTTCCAGAACTATATGCTTCTCTCATTTCTTAGTGCC	884	
Qy	6465	CCCYATGRCNATCTACACTGAACAAGATTTATACAGTTATGTCTATCTTAAGCCCCGCAA	6524	
Db	885	CCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCTATCTTAAGCCCCGCAA	944	
Qy	6525	CAAAAGAGTACCCATCTTCTCTTTGTTATAGGAGCAGGAGTGTCTAGGTGCACCTAGGTAC	6584	
Db	945	CAAAAGAGTACCCATCTTCTCTTTGTTATAGGAGCAGGAGTGTCTAGGTGCACCTAGGTAC	1004	
Qy	6585	TGGCATTTGGGGGTATCAACCTCTACTCAGTTCTTACTTACAAACTATCTCAAGAACTAAA	6644	
Db	1005	TGGCATTTGGGGGTATCAACCTCTACTCAGTTCTTACTTACAAACTATCTCAAGAACTAAA	1064	
Qy	6645	TGGGGACATGGAACGGGTGCGCGACTCGCTGGTCACTTGGCAAGATCAACTTAACCTCCT	6704	
Db	1065	TGGGGACATGGAACGGGTGCGCGACTCGCTGGTCACTTGGCAAGATCAACTTAACCTCCT	1124	
Qy	6705	AGCAGCAGTAGTCTCTTC 6721		

Db 1125 AGCAGCAGTAGTCCTTC 1141

|||||

RESULT 14

US-10-632-793-24

; Sequence 24, Application US/10632793

; Publication No. US20040048298A1

; GENERAL INFORMATION:

; APPLICANT: PARANHOS-BACCALA, Glaucia

; APPLICANT: MALLET, Francois

; APPLICANT: VOISSET, Cecile

; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN

; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT

; FILE REFERENCE: 110048

; CURRENT APPLICATION NUMBER: US/10/632,793

; CURRENT FILING DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: US/09/869,927

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: PCT/FR00/00144

; PRIOR FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: FR 99/00888

; PRIOR FILING DATE: 1999-01-21

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 24

; LENGTH: 1948

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (84)..(84)

; OTHER INFORMATION: n = a o r g o r c o r t/u

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (193)..(193)

; OTHER INFORMATION: n = a o r g o r c o r t/u

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (241)..(241)

; OTHER INFORMATION: n = a o r g o r c o r t/u

US-10-632-793-24

Query Match 9.6%; Score 731; DB 8; Length 1948;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGGAATGACGGTCCCGGAAATATTGATGCCCAATCGTATAGGAGTCT 5744

DB 823 AGAGTTTCTATGAGGAATGACGGTCCCGGAAATATTGATGCCCAATCGTATAGGAGTCT 882

QY 5745 TTSTAAGGGAACCCCGACCTTCACTGCCACACCCCATATGCCCGGCAACTGCTATCACTC 5804

DB 883 TTGTAAGGGAACCCCGACCTTCACTGCCACACCCCATATGCCCGGCAACTGCTATCACTC 942

QY 5805 TGCCACTCTTTGATGATGATGCAATCTCAATTATGGACAGGAAATGATTAATCCTAG 5864

DB 943 TGCCACTCTTTGATGATGCAATCTCAATTATGGACAGGAAATGATTAATCCTAG 1002

QY 5865 TTGTCTCGGAGGACTTGAGTCACTGCTCTTTGGACTTACTTCAACCAACTGCTATGTC 5924

DB 1003 TTGTCTCGGAGGACTTGAGTCACTGCTCTTTGGACTTACTTCAACCAACTGCTATGTC 1062

QY 5925 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCA 5984

DB 1063 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCA 1122

QY 5985 ACTCACCSGGGTACATGGCACTTACGCCCTTCAAGGACTAGATCTCTCAAAACTTACA 6044

DB 1123 ACTCACCSGGGTACATGGCACTTACGCCCTTCAAGGACTAGATCTCTCAAAACTTACA 1182

QY 6045 TGAACACCTCCGTACCCATACTCGCTGGTAAAGCTATTTTAATACCAACCTCACTGGGCT 6104

```

1183 TGAACCCCTCGTACCATCTACTGCGCTGGTAAGCCTATTTAATACACACCCTCACTGGGCT 1242
6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTCAR 6164
1243 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTCAG 1302
6165 GCCATATGTTTCAATCCCTGCTGACCTGAAACCAATGGAACCACTTCAGACACAGAAATAACAC 6224
1303 GCCATATGTTTCAATCCCTGCTGACCTGAAACCAATGGAACCACTTCAGACACAGAAATAACAC 1362
6225 CACTTCGGTTTTAGTAGGACCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAAACCT 6284
1363 CACTTCGGTTTTAGTAGGACCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAAACCT 1422
6285 CACTTCGGTTTTAGTAGGACCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAAACCT 6344
1423 CACTTCGGTTTTAGTAGGACCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAAACCT 1482
6345 AACTCTCCCAACAATAAGTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC 6404
1483 AACTCTCCCAACAATAAGTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC 1542
6405 AGCTATCGTGTGTTGAATGGCTCTTCAGAAATCTATGTGCTTCTCTCAATCTTAGTGCC 6464
1543 AGCTATCGTGTGTTGAATGGCTCTTCAGAAATCTATGTGCTTCTCTCAATCTTAGTGCC 1602
6465 CCCATGRCATCTACACGAAACAGATTTATACAGTTATGTGCTTCTCTCAATCTTAGTGCC 6524
1603 CCTATGCGCATCTACACGAAACAGATTTATACAGTTATGTGCTTCTCTCAATCTTAGTGCC 1662
6525 CAAAGAGTACCAATCTCTCTTTTGTATAGGAGGAGGAGTCTAGGTGCACTAGGTAC 6584
1663 CAAAGAGTACCAATCTCTCTTTTGTATAGGAGGAGGAGTCTAGGTGCACTAGGTAC 1722
6585 TGSCATTGGCGGTATCACAACTCTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAA 6644
1723 TGSCATTGGCGGTATCACAACTCTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAA 1782
6645 TGGGACATGGAACGGGTGCGGACCTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCT 6704
1783 TGGGACATGGAACGGGTGCGGACCTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCT 1842
6705 AGCAGCAGTACCTTC 6721
1843 AGCAGCAGTACCTTC 1859

```

RESULT 15

```

US-10-717-580-5
; Sequence 5, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1948

```

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84)..(84)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (193)..(193)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (241)..(241)
; OTHER INFORMATION: n = any nucleotide
;
US-10-717-580-5

```

```

Query Match          9.6%; Score 731; DB 9; Length 1948;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY	5685	AGAGTTTCTATGAGAAATGACGGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCT	5744
DB	823	AGAGTTTCTATGAGAAATGACGGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCT	882
QY	5745	TTSTAAGGGAACCCGCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC	5804
DB	883	TTGTAAGGGAACCCGCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC	942
QY	5805	TGCGACTCTTTGATGATGCAATCTATTTGAGAGGAAATGATTAATCTCTAG	5864
DB	943	TGCGACTCTTTGATGATGCAATCTATTTGAGAGGAAATGATTAATCTCTAG	1002
QY	5865	TTGTCTCGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCCACCAACTGGTATGTC	5924
DB	1003	TTGTCTCGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCCACCAACTGGTATGTC	1062
QY	5925	TGATGGGGTGGAGTTCAAGATCAGCAAGAGAGAAACATGTAAAGAGAGTAATCTCCCA	5984
DB	1063	TGATGGGGTGGAGTTCAAGATCAGCAAGAGAGAGAAACATGTAAAGAGAGTAATCTCCCA	1122
QY	5985	ACTCACCSGGGTACATGGCACCTCTAGCCCTACAAAGAGCTAGATCTCTCAAACTACA	6044
DB	1123	ACTCACCSGGGTACATGGCACCTCTAGCCCTACAAAGAGCTAGATCTCTCAAACTACA	1182
QY	6045	TGAAACCCCTCGTACCCATCTCGCTGCTGTAAGCCCTATTAAATACCACTCACTGGGCT	6104
DB	1183	TGAAACCCCTCGTACCCATCTCGCTGCTGTAAGCCCTATTAAATACCACTCACTGGGCT	1242
QY	6105	CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTCAG	6164
DB	1243	CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTCAG	1302
QY	6165	GCCATATGTTTCAATCCCTGCTGACCTGAAACCAATGGAACCACTTCAGACAGAAATAACAC	6224
DB	1303	GCCATATGTTTCAATCCCTGCTGACCTGAAACCAATGGAACCACTTCAGACAGAAATAACAC	1362
QY	6225	CACTTCGGTTTTAGTAGGACCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAAACCT	6284
DB	1363	CACTTCGGTTTTAGTAGGACCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAAACCT	1422
QY	6285	CACTTCGTAATAATTTAGCAATATACATACAAACCACTCCCAATGATCAGGTGGGT	6344
DB	1423	CACTTCGTAATAATTTAGCAATATACATACAAACCACTCCCAATGATCAGGTGGGT	1482
QY	6345	AACTCTCCCAACAATAAGTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC	6404
DB	1483	AACTCTCCCAACAATAAGTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC	1542
QY	6405	AGCTATCGTGTGTTGAATGGCTCTTTCAGAAATCTATGTGCTTCTCTCAATCTTAGTGCC	6464
DB	1543	AGCTATCGTGTGTTGAATGGCTCTTTCAGAAATCTATGTGCTTCTCTCAATCTTAGTGCC	1602
QY	6465	CCCATGRCATCTACACGAAACAGATTTATACAGTTATGTGCTTCTCTCAATCTTAGTGCC	6524

```

Db      1603  CCCTATGGCCATCTACATGACAGATTATACAGTTATGTCATATCTAAGCCCGCAA 1662
Qy      6525  CAAAGAGTACCCATCTCTCTTTTGTATAGGACGAGGTGCTAGTGCACTAGGTAC 6584
Db      1663  CAAAGAGTACCCATCTCTCTTTTGTATAGGACGAGGTGCTAGTGCACTAGGTAC 1722
Qy      6585  TGGCAATTGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db      1723  TGGCAATTGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1782
Qy      6645  TGGGACATGGAACGGGTGCGGACTCCTCGTGTACCTTGAAGATCAACTTAACTCCCT 6704
Db      1783  TGGGACATGGAACGGGTGCGGACTCCTCGTGTACCTTGAAGATCAACTTAACTCCCT 1842
Qy      6705  AGCAGCAGTAGTCCTTC 6721
Db      1843  AGCAGCAGTAGTCCTTC 1859

RESULT 16
US-11-028-539-22
; Sequence 22, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/11/028,539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 2055
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2055)
; OTHER INFORMATION:
US-11-028-539-22

Query Match      9.6%; Score 731; DB 13; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      5685  AGAGTTTCTATGGAGATGCAGGTCGCGGAATATTGATGCCCATCGTATAGGAGTCT 5744
Db      495  AGAGTTTCTATGGAGATGCAGGTCGCGGAATATTGATGCCCATCGTATAGGAGTCT 554
Qy      5745  TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTC 5804
Db      555  TTCTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTC 614
Qy      5805  TGCCACTCTTTGCATGCATGCAAAATCACTATTATGGACAGGAAATATGATTAATCCTAG 5864
Db      615  TGCCACTCTTTGCATGCATGCAAAATCACTATTATGGACAGGAAATATGATTAATCCTAG 674
Qy      5865  TTGTCTCGAGACTTGGAGTCACTGTCGTGTTGGACTTACTTCAACCACTGGTATGTC 5924
Db      675  TTGTCTCGAGACTTGGAGTCACTGTCGTGTTGGACTTACTTCAACCACTGGTATGTC 734
Qy      5925  TCATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA 5984
Db      735  TCATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA 794
```

```

Qy      5985  ACTCACCSGGGTACATGSCACCTCTAGCCCTACAAAGAGCTAGATCTCTCAAAACTACA 6044
Db      795  ACTCACCSGGGTACATGSCACCTCTAGCCCTACAAAGAGCTAGATCTCTCAAAACTACA 854
Qy      6045  TGAACCCCTCGTACCCATCTCGCTGGTAAAGCCTATTTAATACCACTCAGTGGGT 6104
Db      855  TGAACCCCTCGTACCCATCTCGCTGGTAAAGCCTATTTAATACCACTCAGTGGGT 914
Qy      6105  CCATGAGGTCTCGGCGCAAAACCCCTACTAACTGTGGATATGCTCCCTCGAATCTCAR 6164
Db      915  CCATGAGGTCTCGGCGCAAAACCCCTACTAACTGTGGATATGCTCCCTCGAATCTCAG 974
Qy      6165  GCCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGCAGAAAATAAACAC 6224
Db      975  GCCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGCAGAAAATAAACAC 1034
Qy      6225  CACTTCCGTTTATAGTAGGACCTCTTGTTCCAATSTGAAATAAACCCATACCTCAAACT 6284
Db      1035  CACTTCCGTTTATAGTAGGACCTCTTGTTCCAATSTGAAATAAACCCATACCTCAAACT 1094
Qy      6285  CACTGTGTAAAAATTTAGCAATATACATACAAACCACTCCCAATGATCAGTGGGT 6344
Db      1095  CACTGTGTAAAAATTTAGCAATATACATACAAACCACTCCCAATGATCAGTGGGT 1154
Qy      6345  AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404
Db      1155  AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1214
Qy      6405  AGCTTATCGTTTGTGAATGGCTCTTACAGAACTATATGTGCTTCTCTCATTTAGTGCC 6464
Db      1215  AGCTTATCGTTTGTGAATGGCTCTTACAGAACTATATGTGCTTCTCTCATTTAGTGCC 1274
Qy      6465  CCCATGRCATCTACACTGAAACAGATTATACAGTTATGTATATCTAAGCCCCGCA 6524
Db      1275  CCCATGRCATCTACACTGAAACAGATTATACAGTTATGTATATCTAAGCCCCGCA 1334
Qy      6525  CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACGAGGTGCTAGTGCACTAGGTAC 6584
Db      1335  CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACGAGGTGCTAGTGCACTAGGTAC 1394
Qy      6585  TGGCATTGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db      1395  TGGCATTGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1454
Qy      6645  TGGGACATGGAACGGGTGCGGACTCCTCGTGTCACTTGAAGATCAACTTAACTCCCT 6704
Db      1455  TGGGACATGGAACGGGTGCGGACTCCTCGTGTCACTTGAAGATCAACTTAACTCCCT 1514
Qy      6705  AGCAGCAGTAGTCCTTC 6721
Db      1515  AGCAGCAGTAGTCCTTC 1531

RESULT 17
US-10-632-793-26
; Sequence 26, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-26

Query Match          9.6%; Score 731; DB 8; Length 2782;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGGAATGAGCGTCCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744
Db      |||||||
QY 867 AGAGTTTCTATGAGGAATGAGCGTCCCGGAATATTGATGCCCATCGTATAGGAGTCT 926
Db      |||||||
QY 5745 TTSTAAGGGAAACCCCACTTCACTGCCCCACATATGCCCCGCAACTGCTATCACTC 5804
Db      |||||||
QY 927 TTCTAAGGGAAACCCCACTTCACTGCCCCACATATGCCCCGCAACTGCTATCACTC 986
Db      |||||||
QY 5805 TGCCACTCTTTGATGATGCAATATCTATTTGAGACAGGAAATGATTAATCCTAG 5864
Db      |||||||
QY 987 TGCCACTCTTTGATGATGCAATATCTATTTGAGACAGGAAATGATTAATCCTAG 1046
Db      |||||||
QY 5865 TTGTCCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGC 5924
Db      |||||||
QY 1047 TTGTCCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGC 1106
Db      |||||||
QY 5925 TGAATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAGAGTATCTCCCA 5984
Db      |||||||
QY 1107 TGAATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAGAGTATCTCCCA 1166
Db      |||||||
QY 5985 ACTCACCSGGGTACATGGCACCCTTAGCCCTTACAAAGAGCTAGATCTCTCAAACTACA 6044
Db      |||||||
QY 1167 ACTCACCSGGGTACATGGCACCCTTAGCCCTTACAAAGAGCTAGATCTCTCAAACTACA 1226
Db      |||||||
QY 6045 TGAACCCCTCCGTAACCTAGCTGCTGTTGGACTTACTTCAACCAACTGGTATGC 6104
Db      |||||||
QY 1227 TGAACCCCTCCGTAACCTAGCTGCTGTTGGACTTACTTCAACCAACTGGTATGC 1286
Db      |||||||
QY 6105 CCATGAGGTCCTGGCCCAAAACCTTACTAACTGTTGGATATGCTTCCCTCACTGGGT 6164
Db      |||||||
QY 1287 CCATGAGGTCCTGGCCCAAAACCTTACTAACTGTTGGATATGCTTCCCTCACTGGGT 1346
Db      |||||||
QY 6165 GGCATATGTTTCAATCCCTGTACCTGAAACAACTTAAAGAGAGTAAATTAACAC 6224
Db      |||||||
QY 1347 GGCATATGTTTCAATCCCTGTACCTGAAACAACTTAAAGAGAGTAAATTAACAC 1406
Db      |||||||
QY 6225 CACTTCCGTTTTAGTAGGACCTCTGTTTCCAATGTTGGAATTAACCCATACCTCAAACT 1466
Db      |||||||
QY 6285 CACTTCCGTTTTAGTAGGACCTCTGTTTCCAATGTTGGAATTAACCCATGATCAGGTGGGT 6344
Db      |||||||
QY 1467 CACTTCCGTTTTAGTAGGACCTCTGTTTCCAATGTTGGAATTAACCCATGATCAGGTGGGT 1526
Db      |||||||
QY 6345 AACTCTCCCAACAATAGTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 6404
Db      |||||||
QY 1527 AACTCTCCCAACAATAGTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 1586
Db      |||||||
QY 6405 AGCCTATGTTGTTGAATGGCTCTTCAGAACTATGTCCTCTCATTTCTTAGTGCC 6464
Db      |||||||
QY 1587 AGCCTATGTTGTTGAATGGCTCTTCAGAACTATGTCCTCTCATTTCTTAGTGCC 1646
Db      |||||||
QY 6465 CCATATGRCATCTACATGMAAAGATTTATACAGTATGTCATATCTAAGCCCGGAA 6524
Db      |||||||
QY 1647 CCTATGACATCTACATGMAAAGATTTATACAGTATGTCATATCTAAGCCCGGAA 1706
Db      |||||||
QY 6525 CAAAAGAGTACCCATCTTCTTTTGTATAGGACAGGAGTGTAGGTGCATAGGTAC 6584
Db      |||||||
QY 1707 CAAAAGAGTACCCATCTTCTTTTGTATAGGACAGGAGTGTAGGTGCATAGGTAC 1766
Db      |||||||
QY 6585 TGGCATTTGGGGTATCAAACTCTTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAA 6644
Db      |||||||
```



```
Db 1407 CACTTCCGTTTGTAGTAGACCTCTTGTTCATGTGGAATTAACCCATACCTCAACCT 1466
Qy 6285 CACTGTGTAAATTTAGCAATPACTACATACACAAACCACTCCCAATGCATCAGGTGGGT 6344
Db 1467 CACCTGTGTAAATTTAGCAATPACTACATACACAAACCACTCCCAATGCATCAGGTGGGT 1526
Qy 6345 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404
Db 1527 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1586
Qy 6405 AGCTATCGTTGTTGTAATGGCTCTTCAAGATCTATGTGCTCTCATTTCTAGTGCC 6464
Db 1587 AGCTATCGTTGTTGTAATGGCTCTTCAAGATCTATGTGCTCTCATTTCTAGTGCC 1646
Qy 6465 CCCATGRCATCTACCTGAAACAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 6524
Db 1647 CCCTATGACCATCTACCTGAAACAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 1706
Qy 6525 CAAAAGAGTACCATCTCTCTTTTGTATAGGACAGAGTGCTAGGTGCATAGGTAC 6584
Db 1707 CAAAAGAGTACCATCTCTCTTTTGTATAGGACAGAGTGCTAGGTGCATAGGTAC 1766
Qy 6585 TGGCAATGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1767 TGGCAATGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1826
Qy 6645 TGGGACATGGAACGGGTGCGGACTCCCTGCTGCTACCTTTGCAAGATCAACTTTAATCCCT 6704
Db 1827 TGGGACATGGAACGGGTGCGGACTCCCTGCTGCTACCTTTGCAAGATCAACTTTAATCCCT 1886
Qy 6705 AGCAGAGTAGTCTTTC 6721
Db 1887 AGCAGAGTAGTCTTTC 1903

RESULT 19
US-10-717-580-7
; Sequence 7, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Hervé
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Human
US-10-717-580-7
```

```
Query Match 9.6%; Score 731; DB 9; Length 2782;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5685 AGAGTTTCTATGAGAAATCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 867 AGAGTTTCTATGAGAAATCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 926
```

```
Qy 5745 TTSTAAGGGAACCCACCTTCACTGCCACACCATATGCCGCACTGCTATCACTC 5804
Db 927 TTCTAAGGGAACCCACCTTCACTGCCACACCATATGCCGCACTGCTATCACTC 986
Qy 5805 TGCCACATCTTTGCAATGCAATATCTATTTGGACAGGAAATATGATTAATCCTAG 5864
Db 987 TGCCACATCTTTGCAATGCAATATCTATTTGGACAGGAAATATGATTAATCCTAG 1046
Qy 5865 TTGTCCTGGAGGACTTGGAGTCACTGTCTGTGGAATTTACCCCAAACTGGTATGTC 5924
Db 1047 TTGTCCTGGAGGACTTGGAGTCACTGTCTGTGGAATTTACCCCAAACTGGTATGTC 1106
Qy 5925 TGATGGGGGTGGAGTTCAGATCAGGCAAGAAACATGTGTAAGAGTAATCTCCCA 5984
Db 1107 TGATGGGGGTGGAGTTCAGATCAGGCAAGAAACATGTGTAAGAGTAATCTCCCA 1166
Qy 5985 ACTCACSGGGGTACATGSCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044
Db 1167 ACTCACSGGGGTACATGSCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 1226
Qy 6045 TGAACCCCTCCGTACCCATCTCGCTGGTAAGGCTATTTAATACCACTCCTGAGGT 6104
Db 1227 TGAACCCCTCCGTACCCATCTCGCTGGTAAGGCTATTTAATACCACTCCTGAGGT 1286
Qy 6105 CCATGAGTCTCGGCGCCAAAACCTCTAACTGTTGGATATGCCCTCCCTGAACTTCAR 6164
Db 1287 CCATGAGTCTCGGCGCCAAAACCTCTAACTGTTGGATATGCCCTCCCTGAACTTCAR 1346
Qy 6165 GCCATATGTTTCAATCCCTGTACCTGAAACAACTTTGAGCAAGAAATAAACAAC 6224
Db 1347 GCCATATGTTTCAATCCCTGTACCTGAAACAACTTTGAGCAAGAAATAAACAAC 1406
Qy 6225 CACTTCCGTTTGTAGTAGGACCTCTGTTTCAATSTGGAATTAACCCATACCTCAAACT 6284
Db 1407 CACTTCCGTTTGTAGTAGGACCTCTGTTTCAATSTGGAATTAACCCATACCTCAAACT 1466
Qy 6285 CACTTGTGTAATTTAGCAATACATACACAAACCACTCCCAATGCATCAGGTGGGT 6344
Db 1467 CACTTGTGTAATTTAGCAATACATACACAAACCACTCCCAATGCATCAGGTGGGT 1526
Qy 6345 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTAACCTC 6404
Db 1527 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTAACCTC 1586
Qy 6405 AGCTATCGTTTGTGTAATGGCTCTTCAAGATCTATGTGCTCTCATTTCTTAGTGCC 6464
Db 1587 AGCTATCGTTTGTGTAATGGCTCTTCAAGATCTATGTGCTCTCATTTCTTAGTGCC 1646
Qy 6465 CCCATGRCATCTACCTGAAACAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 6524
Db 1647 CCCTATGACCATCTACCTGAAACAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 1706
Qy 6525 CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACAGAGTGCTAGGTGCATAGGTAC 6584
Db 1707 CAAAAGAGTACCCATCTTCTCTTTTGTATAGGACAGAGTGCTAGGTGCATAGGTAC 1766
Qy 6585 TGGCAATGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1767 TGGCAATGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1826
Qy 6645 TGGGACATGGAACGGGTGCGGACTCCCTGCTGCTACCTTTGCAAGATCAACTTTAATCCCT 6704
Db 1827 TGGGACATGGAACGGGTGCGGACTCCCTGCTGCTACCTTTGCAAGATCAACTTTAATCCCT 1886
Qy 6705 AGCAGAGTAGTCTTTC 6721
Db 1887 AGCAGAGTAGTCTTTC 1903
```

```
RESULT 20
US-09-902-535-1
; Sequence 1, Application US/09902535
```



```
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: MI, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; and treating preclampsia and gestational trophoblast
; FILE OF INVENTION: disorders
; FILE REFERENCE: GIN-6008B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930)...(2546)
; US-09-902-535-1

Query Match          9.6%; Score 728; DB 3; Length 2930;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1898; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 4802 AGCCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAAA 4861
DB 152 AGCCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAAA 211
QY 4862 TAATTCCTGAGTTCGACCTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGG 4921
DB 212 TAATTCCTGAGTTCGACCTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGG 271
QY 4922 CCACAGTAACCCAGGAGTATCCAGGCTTAGGTATACGATATCATCTACAGTCCGCT 4981
DB 272 CCACAGTAACCCAGGAGTATCCAGGCTTAGGTATACGATATCATCTACAGTCCGCT 331
QY 4982 GAAGGCCACAGTCTCAGGAGAGGTGAGAAAATGAATGAATACTCAAGAGGACATCTAA 5041
DB 332 GAAGGCCACAGTCTCAGGAGAGGTGAGAAAATGAATGAATACTCAAGAGGACATCTAA 391
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTGGCTTATAGCCCTTAAAAA 5101
DB 392 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGTTCTGTGGCTTATAGCCCTTAAAAA 451
QY 5102 GAATCTGCAACTTTCCTCCCAAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
DB 452 GAATCTGCAACTTTCCTCCCAAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 511
QY 5162 CCTTCATAACCAATGACCTTGTGCTTGACCCAGAGCAGCCACTTAGTTGCGACATCAC 5221
DB 512 CCTTCATAACCAATGACCTTGTGCTTGACCCAGAGCAGCCACTTAGTTGCGACATCAC 571
QY 5222 CTCCTTAGCCAAATATCAACAAGTCTTAAAAACATTAACAGGAACCTATCCCTCGAGAAGA 5281
DB 572 CTCCTTAGCCAAATATCAACAAGTCTTAAAAACATTAACAGGAACCTATCCCTCGAGAAGA 631
QY 5282 GGGAAAAGAACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTTAATTC 5341
DB 632 GGGAAAAGAACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTTAATTC 691
QY 5342 CCCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCAATTTATYATACCCCACTGC 5401
DB 692 CCCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCAATTTATYATACCCCACTGC 751
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACCTTGGATCAAAATCTCTGGATCT 5461
DB 752 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACCTTGGATCAAAATCTCTGGATCT 811
QY 5462 GCCAAAGGAAACCTGAAAAATCCAGGAGCAACGCTAGCTATTCTGTGAAACCTCTAGAGGA 5521
```

```
DB 812 GCCAAAGGAAACCTGAAAAATCCAGGAGCAACGCTAGCTATTCTGTGAATCTTAGAGGA 871
QY 5522 TTTGCGCCTGCTCTTCAAAACAACAACAGGAGAAAGTAACTAAATCATAAATCCCCCA 5581
DB 872 TTTGCGCCTGCTCTTCAAAACAACAACAGGAGAAAGTAACTAAATCATAAATCCCCCA 930
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCCTCTTTCACCTCTCATCTG 5641
DB 931 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCCTCTTTCACCTCTCATCTG 990
QY 5642 CACCCCTCCATGCGCCCTGTATGACAGTACCTCCCTTACCAAGAGTCTTCTATGAGAA 5701
DB 991 CACCCCTCCATGCGCCCTGTATGACAGTACCTCCCTTACCAAGAGTCTTCTATGAGAA 1050
QY 5702 TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCCA 5761
DB 1051 TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCA 1110
QY 5762 CTTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 5821
DB 1111 CTTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 1170
QY 5822 ATGCAATATCTCATTTATGGAAGGAAATGATTAATCTAGTCTTCTGAGGAGTCTG 5881
DB 1171 ATGCAATATCTCATTTATGGAAGGAAATGATTAATCTAGTCTTCTGAGGAGTCTG 1230
QY 5882 GAGTCACTGCTGTGAGACTTCTTCAACCAACTGCTATGCTGATGCGGGTGGAGTTC 5941
DB 1231 GAGTCACTGCTGTGAGACTTCTTCAACCAACTGCTATGCTGATGCGGGTGGAGTTC 1290
QY 5942 AAGATCAGGCAAGAGAAAAACATGTAAAGAGAGTAACTCCCAACTCACCGGGTACATG 6001
DB 1291 AAGATCAGGCAAGAGAAAAACATGTAAAGAGAGTAACTCCCAACTCACCGGGTACATG 1350
QY 6002 GCACCTCTAGCCCTTCAAAAGGACTAGATCTCTCAAACTCATATGAAACCCCTCCGTACCC 6061
DB 1351 GCACCTCTAGCCCTTCAAAAGGACTAGATCTCTCAAACTCATATGAAACCCCTCCGTACCC 1410
QY 6062 ATACTCGCTGCTGTAAGCTTATTAATACCACCTCACTGGGCTCATAGGCTCTCGGCC 6121
DB 1411 ATACTCGCTGCTGTAAGCTTATTAATACCACCTCACTGGGCTCATAGGCTCTCGGCC 1470
QY 6122 AAAACCCCTAACTGTTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 6181
DB 1471 AAAACCCCTAACTGTTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 1530
QY 6182 CTGTACTGAAACAATGGAAACAACCTTCAGCAGCAAGAAATAAACCACTTCGGTTTATGATG 6241
DB 1531 CTGTACTGAAACAATGGAAACAACCTTCAGCAGCAAGAAATAAACCACTTCGGTTTATGATG 1590
QY 6242 GACCTCTGTTTCCCAATSTGGAAATAAACCCATACCTCAACCTCACTCTGTGTAATAATTA 6301
DB 1591 GACCTCTGTTTCCCAATSTGGAAATAAACCCATACCTCAACCTCACTCTGTGTAATAATTA 1650
QY 6302 GCAATACTACATACACCAACCACTCCCAATGATCAGGTGGGTAACTCTCCACACAA 6361
DB 1651 GCAATACTACATACACCAACCACTCCCAATGATCAGGTGGGTAACTCTCCACACAA 1710
QY 6362 TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCTGTTTGA 6421
DB 1711 TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCTGTTTGA 1770
QY 6422 ATGGCTCTTCAGAACTATGCTCTCTCTCATCTTAGTGGCCCTCATGRCATCTACA 6481
DB 1771 ATGGCTCTTCAGAACTATGCTCTCTCTCATCTTAGTGGCCCTCATGRCATCTACA 1830
QY 6482 CTGAAACAAGATTATACAGTATGATCATATCTAAGCCCGCAACAAAAGAGTACCCATTC 6541
DB 1831 CTGAAACAAGATTATACAGTATGATCATATCTAAGCCCGCAACAAAAGAGTACCCATTC 1890
QY 6542 TTTCTTTTGTATAGGAGCAGGAGTGTAGGTGCACTAGGTACTGTGCGGTATCA 6601
```

Db 1891 TTCTTTTGTATTAGAGCAGGAGTGCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATCA 1950  
Qy 6602 CAACTCTTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 6661  
Db 1951 CAACTCTTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 2010  
Qy 6662 TGGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCCTTC 6721  
Db 2011 TGGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCCTTC 2070

RESULT 21  
US-10-632-793-21  
; Sequence 21, Application US/10632793  
; Publication No. US20040048298A1  
; GENERAL INFORMATION:  
; APPLICANT: PARANHOS-BACCALA, Glaucia  
; APPLICANT: MALLET, Francois  
; APPLICANT: VOISSET, Cecile  
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN  
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT  
; FILE REFERENCE: 110048  
; CURRENT APPLICATION NUMBER: US/10/632,793  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/869,927  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: PCT/FR00/00144  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: FR 99/00888  
; PRIOR FILING DATE: 1999-01-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 2938  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-632-793-21

Query Match 9.5%; Score 724; DB 8; Length 2938;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1476 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTGCTTC 1535  
Db 883 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTGCTTC 942

Qy 1536 TGGAGAGACTAAGGGAGGCAATTGAGGAAGCGTCTCTGTCACCTGACTCTTCTGAAG 1595  
Db 943 TGGAGAGACTAAGGGAGGCAATTGAGGAAGCGTCTCTGTCACCTGACTCTTCTGAAG 1002

Qy 1596 GCCAACTAATCTTAAAGCGTAAGTTTATCAGTCAGTCAGCTGCAGACATTAGAAAAAAT 1655  
Db 1003 GCCAACTAATCTTAAAGCGTAAGTTTATCAGTCAGTCAGCTGCAGACATTAGAAAAAAT 1062

Qy 1656 TCAAAAGTCTGCGTAGGCGCGGAGCAAACTTAGAAAACTTATGAACTTGGCACTTC 1715  
Db 1063 TCAAAAGTCTGCGTAGGCGCGGAGCAAACTTAGAAAACTTATGAACTTGGCACTTC 1122

Qy 1716 GGTTTTTTATAATAGAGATCAGGAGGAGCGGAAACAGGACAAACGGGATTTAAAAAAA 1775  
Db 1123 GGTTTTTTATAATAGAGATCAGGAGGAGCGGAAACAGGACAAACGGGATTTAAAAAAA 1182

Qy 1776 AGGCCACCGCTTAGTCATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTGAAAAAGGA 1835  
Db 1183 AGGCCACCGCTTAGTCATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTGAAAAAGGA 1242

Qy 1836 AAGCTGGGCAATTAAGTCCCTAATAGGCTTGTTCAGTGCAGTCCGCTTACAGGACACT 1895  
Db 1243 AAGCTGGGCAATTAAGTCCCTAATAGGCTTGTTCAGTGCAGTCCGCTTACAGGACACT 1302

Qy 1896 TTAATAAAGATGTCACAGTAGAAGTAAGCGGCCCTTCGTCCATGCCCTTTATTTCAAG 1955  
Db 1303 TTAATAAAGATGTCACAGTAGAAGTAAGCGGCCCTTCGTCCATGCCCTTTATTTCAAG 1362

Qy 1956 GGAATCACTGGAAGSCCAGTCCCTGAGGGGACAAAAGCTCTTTTGAATCAGAAGCCACTA 2015  
Db 1363 GGAATCACTGGAAGSCCAGTCCCTGAGGGGACAAAAGTCTTTTGAATCAGAAGCCACTA 1422

Qy 2016 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCGATCA 2075  
Db 1423 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCGATCA 1482

Qy 2076 CCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAGGTTGTCTCTGGACA 2135  
Db 1483 CCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAGGTTGTCTCTGGACA 1542

Qy 2136 CTGGTGGGCTCTCTTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2195  
Db 1543 CTGGTGGGCTCTCTTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1602

Qy 2196 CTAT 2199  
Db 1603 CTAT 1606

RESULT 22  
US-10-717-580-2  
; Sequence 2, Application US/10717580  
; Publication No. US20040176314A1  
; GENERAL INFORMATION:  
; APPLICANT: BESEME, Frederic  
; APPLICANT: BLOND, Jean-Luc  
; APPLICANT: BOUTON, Olivier  
; APPLICANT: MANDRAND, Bernard  
; APPLICANT: MALLET, Francois  
; APPLICANT: PERRON, Herve  
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASE  
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS  
; FILE REFERENCE: 105045  
; CURRENT APPLICATION NUMBER: US/10/717,580  
; CURRENT FILING DATE: 2003-11-21  
; PRIOR APPLICATION NUMBER: US/09/446,024A  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/FR98/01442  
; PRIOR FILING DATE: 1998-07-06  
; PRIOR APPLICATION NUMBER: FR 97/08815  
; PRIOR FILING DATE: 1997-07-07  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2938  
; TYPE: DNA  
; ORGANISM: Human  
US-10-717-580-2

Query Match 9.5%; Score 724; DB 9; Length 2938;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1476 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTGCTTC 1535  
Db 883 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTGCTTC 942

Qy 1536 TGGAGAGACTAAGGGAGGCAATTGAGGAAGCGTCTCTGTCACCTGACTCTTCTGAAG 1595  
Db 943 TGGAGAGACTAAGGGAGGCAATTGAGGAAGCGTCTCTGTCACCTGACTCTTCTGAAG 1002

Qy 1596 GCCAACTAATCTTAAAGCGTAAGTTTATCAGTCAGTCAGCTGCAGACATTAGAAAAAAT 1655  
Db 1003 GCCAACTAATCTTAAAGCGTAAGTTTATCAGTCAGTCAGCTGCAGACATTAGAAAAAAT 1062

Qy 1656 TCAAAAGTCTGCGTAGGCGCGGAGCAAACTTAGAAAACTTATGAACTTGGCACTTC 1715  
Db 1063 TCAAAAGTCTGCGTAGGCGCGGAGCAAACTTAGAAAACTTATGAACTTGGCACTTC 1122

Qy 1716 GGTTTTTTATAATAGAGATCAGGAGGAGCGGAAACAGGACAAACGGGATTTAAAAAAA 1775



```
; Sequence 3, Application US/10016249
; Publication No. US200301000531
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/10/016,249
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US/09/175,928
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-249-3

Query Match      9.0%; Score 680; DB 6; Length 2946;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGAAATGACGGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744
Db 1032 AGAGTTTCTATGGAGAAATGACGGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 1091
QY 5745 TTSTAAGGGAAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804
Db 1092 TTCTAAGGGAAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 1151
QY 5805 TGCCACTCTTTGCATGCATGCAAAATCACTATTATTTGGACAGAAATAATGATTAATCTTAG 5864
Db 1152 TGCCACTCTTTGCATGCATGCAAAATCACTATTATTTGGACAGAAATAATGATTAATCTTAG 1211
QY 5865 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAAACTGGTATGTC 5924
Db 1212 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAAACTGGTATGTC 1271
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAATGATCTCCCA 5984
Db 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAATGATCTCCCA 1331
QY 5985 ACTCACCGGGTACATGGACCTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTACA 6044
Db 1332 ACTCACCGGGTACATGGACCTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTACA 1391
QY 6045 TGAACCCCTCCGTACCCATCTCGCTGTAGCCCTTTATATACACCCCTCACTGGGCT 6104
Db 1392 TGAACCCCTCCGTACCCATCTCGCTGTAGCCCTTTATATACACCCCTCACTGGGCT 1451
QY 6105 CCATGAGGTCTGGGCCAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCAR 6164
Db 1452 CCATGAGGTCTGGGCCAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCAG 1511
QY 6165 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAAACAATTTAGACAGAAATAAACAC 6224
Db 1512 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAAACAATTTAGACAGAAATAAACAC 1571
QY 6225 CACTTCCGTTTTAGTAGGACCTCTGTTCCTCAATSTGGAAATAACCCATCTCAAACT 6284
Db 1572 CACTTCCGTTTTAGTAGGACCTCTGTGTTCCTCAATCTGGAATAAACCCATCTCAAACT 1631
QY 6285 CACTGTGTAAATTTAGCAATACATACACCAACTCCCAATGATCAGGTGGGT 6344
Db 1632 CACTGTGTAAATTTAGCAATACATACACCAACTCCCAATGATCAGGTGGGT 1691

; Sequence 543, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 543
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-543

Query Match      8.6%; Score 651; DB 10; Length 3633;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGAAATGACGGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744
Db 866 AGAGTTTCTATGGAGAAATGACGGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 925
QY 5745 TTSTAAGGGAAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804
Db 926 TTCTAAGGGAAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 985
QY 5805 TGCCACTCTTTGCATGCATGCAAAATCACTATTATTTGGACAGAAATAATGATTAATCTTAG 5864
Db 986 TGCCACTCTTTGCATGCATGCAAAATCACTATTATTTGGACAGAAATAATGATTAATCTTAG 1045
QY 5865 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAAACTGGTATGTC 5924
Db 1046 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAAACTGGTATGTC 1105

RESULT 25
US-10-821-234-543
; Sequence 543, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 543
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-543
```

```
QY 5925 TGATGGGGTGGAGCTTCAAGATCAGGCAAGAGAAAACATGTAAGAAAGTAATCTCCCA 5984
Db 1106 TGATGGGGTGGAGCTTCAAGATCAGGCAAGAGAAAACATGTAAGAAAGTAATCTCCCA 1165
QY 5985 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 6044
Db 1166 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 1225
QY 6045 TGAACCTTCGGTACCATCTAGCCCTGGTAAAGCTATTTAATAACCAACCTCTAGTGGCT 6104
Db 1226 TGAACCTTCGGTACCATCTAGCCCTGGTAAAGCTATTTAATAACCAACCTCTAGTGGCT 1285
QY 6105 CCATGAGTCTCGGCCCAAAACCTTACTACTGTTGGATATGCTCCCTGCTGACTTCAR 6164
Db 1286 CCATGAGTCTCGGCCCAAAACCTTACTACTGTTGGATATGCTCCCTGCTGACTTCAG 1345
QY 6165 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTACAGCACAGAAATAAACAC 6224
Db 1346 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTACAGCACAGAAATAAACAC 1405
QY 6225 CACTTCGGTTTTAGTAGGACTCTGTTGTTTCCAATSTGGAATAAACCCATACCTCAAACT 6284
Db 1406 CACTTCGGTTTTAGTAGGACTCTGTTGTTTCCAATSTGGAATAAACCCATACCTCAAACT 1465
QY 6285 CACTGCTGTAATAATTAGCAATACTACATACACAACTCCCAATGCAATGCACTAGTGGT 6344
Db 1466 CACTGCTGTAATAATTAGCAATACTACATACACAACTCCCAATGCAATGCACTAGTGGT 1525
QY 6345 AACTCTCCCAACAAATAGTCTGCTTACCTCAGGAATATTTTCTGCTGCTGCTC 6404
Db 1526 AACTCTCCCAACAAATAGTCTGCTTACCTCAGGAATATTTTCTGCTGCTGCTC 1585
QY 6405 AGCTATGCTGTTGTAATGGCTCTTCAGAACTATGCTGCTCTCTCATCTTAGTGCC 6464
Db 1586 AGCTATGCTGTTGTAATGGCTCTTCAGAACTATGCTGCTCTCTCATCTTAGTGCC 1645
QY 6465 CCCATGRCATCTACACTGAAAGATTTATACAGTTATGCTATCTAAGCCCGCAA 6524
Db 1646 CCTATGACCATCTACACTGAAAGATTTATACAGTTATGCTATCTAAGCCCGCAA 1705
QY 6525 CAAAGAGTACCACTTCTCTTTTGTATAGAGCAGGAGTGTAGTGCATAGGTAC 6584
Db 1706 CAAAGAGTACCACTTCTCTTTTGTATAGAGCAGGAGTGTAGTGCATAGGTAC 1765
QY 6585 TGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
Db 1766 TGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 1825
QY 6645 TGGGACATGGAAGGGTCCCGACTCCCTGGTCACTTCCAGATCA 6692
Db 1826 TGGGACATGGAAGGGTCCCGACTCCCTGGTCACTTCCAGATCA 1873
```

## RESULT 26

US-10-632-793-23

```
; Sequence 23, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucaia
; APPLICANT: MALLET, Françoise
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
```

## RESULT 27

US-10-717-580-4

```
; Sequence 4, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERON, Hervé
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DIS
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
```

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 23

; LENGTH: 2006

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (305)..(305)

; OTHER INFORMATION: n = a or g or c or t/u

US-10-632-793-23

Query Match

Best Local Similarity 5.3%; Score 399; DB 8; Length 2006;

Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 4802 AGSCCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 4861
Db 643 AGSCCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 702
QY 4862 TAATTCAGATTCGGACTTCCCAGAGGTTTACAGAGTGAACAATAGCCCTGCTTCCAGG 4921
Db 703 TAATTCAGATTCGGACTTCCCAGAGGTTTACAGAGTGAACAATAGCCCTGCTTCCAGG 762
QY 4922 CCACAGTAACCCAGGAGGATATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 4981
Db 763 CCACAGTAACCCAGGAGGATATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 822
QY 4982 GAAGGCCACAGTCTCAGGGAAGGTCGAGAAAATGAATAAAGTGAATAAAGGACATCTAA 5041
Db 823 GAAGGCCACAGTCTCAGGGAAGGTCGAGAAAATGAATAAAGTGAATAAAGGACATCTAA 882
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTTATAGCCTTAAAAA 5101
Db 883 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTTATAGCCTTAAAAA 942
QY 5102 GAATCTGCAACTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
Db 943 GAATCTGCAACTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 1002
QY 5162 CCTTCATAACCAATGACCTTGTGTCACCAAGAGCAGCCACTTGTAGTGCACATCAC 5221
Db 1003 CCTTCATAACCAATGACCTTGTGTCACCAAGAGCAGCCACTTGTAGTGCACATCAC 1062
QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTTACAGGAACCTATCCCTGAGAAGA 5281
Db 1063 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTTACAGGAACCTATCCCTGAGAAGA 1122
QY 5282 GGGAAAAGAACTATTCCACC 5302
Db 1123 GGGAAAAGAACTATTCCACC 1143
```



```

; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-28

Query Match      5.0%; Score 380; DB 8; Length 3372;
Best Local Similarity 99.6%; Pred. No. 1.2e-190;
Matches 480; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4802 AGGCTTCCCTGTAGACAGAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 4861
Db      |
QY 2032 AGGCTTCCCTGTAGACAGAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 2091
Db      |
QY 4862 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCAGG 4921
Db      |
QY 2092 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCAGG 2151
Db      |
QY 4922 CCACAGTAACCCAGGAGTATCCCGAGGCTTACGATATACGATATCACTTACACTGGCCCT 4981
Db      |
QY 2152 CCACAGTAACCCAGGAGTATCCCGAGGCTTACGATATACGATATCACTTACACTGGCCCT 2211
Db      |
QY 4982 GAAGGCCACAGTCTCAGGAGGTCGAGAAAGTGAATGAATAAAGGACATCTAA 5041
Db      |
QY 2212 GAAGGCCACAGTCTCAGGAGGTCGAGAAAGTGAATGAATAAAGGACATCTAA 2271
Db      |
QY 5042 AAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGCTTATAGCCTTAAAAA 5101
Db      |
QY 2272 AAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGCTTATAGCCTTAAAAA 2331
Db      |
QY 5102 GAATCTGCAACTTTCCTCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
Db      |
QY 2332 GAATCTGCAACTTTCCTCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 2391
Db      |
QY 5162 CCTTCATAACCAATGACCTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 5221
Db      |
QY 2392 CCTTCATAACCAATGACCTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 2451
Db      |
QY 5222 CTCCTTAGCCAAATATCAAGTTCCTTAAAAACATTACAAGGAACCTTATCCCTGAGAAGA 5281
Db      |
QY 2452 CTCCTTAGCCAAATATCAAGTTCCTTAAAAACATTACAAGGAACCTTATCCCTGAGAAGA 5281
Db      |
QY 5282 GG 5283
Db      ||
QY 2512 GG 2513
Db      ||

RESULT 30
US-10-717-580-9
; Sequence 9, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISE
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 3372
; TYPE: DNA

```

```

; ORGANISM: Human
US-10-717-580-9

Query Match      5.0%; Score 380; DB 9; Length 3372;
Best Local Similarity 99.6%; Pred. No. 1.2e-190;
Matches 480; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4802 AGGCTTCCCTGTAGACAGAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 4861
Db      |
QY 2032 AGGCTTCCCTGTAGACAGAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 2091
Db      |
QY 4862 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCAGG 4921
Db      |
QY 2092 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCAGG 2151
Db      |
QY 4922 CCACAGTAACCCAGGAGTATCCCGAGGCTTACGATATACGATATCACTTACACTGGCCCT 4981
Db      |
QY 2152 CCACAGTAACCCAGGAGTATCCCGAGGCTTACGATATACGATATCACTTACACTGGCCCT 2211
Db      |
QY 4982 GAAGGCCACAGTCTCAGGAGGTCGAGAAAGTGAATGAATAAAGGACATCTAA 5041
Db      |
QY 2212 GAAGGCCACAGTCTCAGGAGGTCGAGAAAGTGAATGAATAAAGGACATCTAA 2271
Db      |
QY 5042 AAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGCTTATAGCCTTAAAAA 5101
Db      |
QY 2272 AAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGCTTATAGCCTTAAAAA 2331
Db      |
QY 5102 GAATCTGCAACTTTCCTCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
Db      |
QY 2332 GAATCTGCAACTTTCCTCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 2391
Db      |
QY 5162 CCTTCATAACCAATGACCTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 5221
Db      |
QY 2392 CCTTCATAACCAATGACCTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 2451
Db      |
QY 5222 CTCCTTAGCCAAATATCAAGTTCCTTAAAAACATTACAAGGAACCTTATCCCTGAGAAGA 5281
Db      |
QY 2452 CTCCTTAGCCAAATATCAAGTTCCTTAAAAACATTACAAGGAACCTTATCCCTGAGAAGA 5281
Db      |
QY 5282 GG 5283
Db      ||
QY 2512 GG 2513
Db      ||

Search completed: August 2, 2006, 15:51:59
Job time : 7491 secs

```

**This Page Blank (uspto)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:38 ; Search time 4018 Seconds

(without alignments)  
13156.692 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caacaatcggtataataacc.....tattaaatctgcartctgcr 7582

Scoring table:

OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 5244920 seqs, 3486124231 residues

Word size : 6

Total number of hits satisfying chosen parameters: 10438277

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : N Geneseq 8.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*
- 15: Geneseqn2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7582	100.0	7582	3	Aa559215 Human end
2	7378	97.3	7582	2	Aa552565 Complete
3	979	12.9	7466	5	Aa568626 DNA encod
4	888	11.7	2372	2	Aa525664 Human end
5	888	11.7	2372	3	Aa559214
6	779	10.3	2599	3	Abn97927 Human ret
7	779	10.3	8523	7	Ady30988 Human gen
8	779	10.3	8523	7	Ady36376 HIRA geno
9	779	10.3	10499	3	Abn97929 Human ret
10	779	10.3	56093	6	Abn97929 Human ret
11	750	9.9	6394	5	Aa584210 Colon ade
12	731	9.6	1614	14	AED10984
13	731	9.6	1614	14	AED10986
14	731	9.6	1617	5	AaH20070 HERV-R393
15	731	9.6	1617	14	AED10942
16	731	9.6	1948	2	Aa525659 HERV enve
17	731	9.6	1948	2	Aa525659 Human end
18	731	9.6	2055	3	Abn97948 Human ret

19	731	9.6	2782	2	AaX25661	AaX25661 Human end
20	731	9.6	2782	3	AaA59211	AaA59211 5' non co
21	731	9.6	2782	5	AaH20069	AaH20069 HERV-W en
22	728	9.6	2930	6	AaD24195	AaD24195 Human syn
23	724	9.5	2938	2	AaX25656	AaX25656 Human end
24	724	9.5	2938	3	AaA59206	AaA59206 Gag and p
25	699	9.2	2902	10	ADP59718	ADP59718 Human con
26	680	9.0	2946	2	AaX77526	AaX77526 Human sec
27	680	9.0	2946	3	AaA59468	AaA59468 Human sec
28	680	9.0	2946	10	ADC38776	ADC38776 Human cDN
29	680	9.0	6332	13	ADU81027	ADU81027 Membrane-
30	651	8.6	2781	5	AaF55630	AaF55630 Nucleotid
31	651	8.6	3633	14	AED73715	AED73715 Human pla
32	339	5.3	2006	2	AaX25658	AaX25658 Human end
33	339	5.3	2006	3	AaA59208	AaA59208 3' pol ge
34	339	5.3	3464	10	ADE09587	ADE09587 Novel DNA
35	339	5.3	8294	5	AaA584209	AaA584209 DNA encod
36	339	5.3	8294	10	ADP59732	ADP59732 Human con
37	380	5.0	3372	2	AaX25663	AaX25663 Human end
38	380	5.0	3372	3	AaA59213	AaA59213 Partial p
39	380	5.0	3831	5	AaA59213	AaA59213 DNA encod
40	380	5.0	5154	5	AaA567609	AaA567609 DNA encod
41	380	5.0	8279	5	AaA576474	AaA576474 DNA encod
42	362	4.8	4349	5	AaA576475	AaA576475 DNA encod
43	267	3.5	1326	3	ABN97928	ABN97928 Human ret
44	249	3.3	1136	2	AaX25660	AaX25660 Human end
45	249	3.3	1136	3	AaA59210	AaA59210 3' pol ge
46	213	2.8	1321	2	AaX25655	AaX25655 Human end
47	213	2.8	1321	3	AaA59205	AaA59205 5' non co
48	213	2.8	1715	5	AaA573701	AaA573701 DNA encod
49	211	2.6	2264	15	AEF75131	AEF75131 Human ret
50	194	2.6	2264	15	AEF75131	AEF75131 Human pol
51	190	2.5	387	5	AaA584202	AaA584202 DNA encod
52	190	2.5	542	2	AaA59212	AaA59212 3' pol ge
53	190	2.5	542	2	AAV15134	AAV15134 Human adu
54	183	2.4	1080	3	ABN97949	ABN97949 Human ret
55	164	2.2	1021	10	ADE07345	ADE07345 Novel cod
56	159	2.1	384	2	AAV86331	AAV86331 EST clone
57	158	2.1	666	2	AaX25662	AaX25662 Human end
58	158	2.1	666	3	AaA59212	AaA59212 3' pol ge
59	157	2.1	910	5	AaA571725	AaA571725 DNA encod
60	149	2.0	443	3	AaC17319	AaC17319 Human sec
61	138	1.8	783	2	AaX25669	AaX25669 Human end
62	123	1.6	549	14	ACL55948	ACL55948 Human col
63	104	1.4	893	5	AaA565964	AaA565964 DNA encod
64	89	1.2	711	3	ABN97947	ABN97947 Human ret
65	85	1.1	2074	6	AD41225	AD41225 Human EMB
66	76	1.0	46340	3	ABN97978	ABN97978 Human ret
67	73	1.0	902	5	AaA571722	AaA571722 DNA encod
68	73	1.0	1039	5	AaA584206	AaA584206 DNA encod
69	73	1.0	1341	5	AaA584194	AaA584194 DNA encod
70	73	1.0	1341	5	AaA571719	AaA571719 DNA encod
71	73	1.0	1481	2	AaA43217	AaA43217 Multiple
72	73	1.0	1481	2	AaX29703	AaX29703 Clone C15
73	73	1.0	1481	12	ADG14846	ADG14846 MSRV asso
74	73	1.0	1629	3	AaA96625	AaA96625 DNA encod
75	73	1.0	2030	3	AaA63826	AaA63826 Nucleotid
76	72	0.9	575	5	AaA572226	AaA572226 DNA encod
77	72	0.9	2049	5	AaA590991	AaA590991 DNA encod
78	72	0.9	4805	5	AaA573910	AaA573910 DNA encod
79	72	0.9	4805	5	AaA573604	AaA573604 DNA encod
80	72	0.9	4805	5	AaA584191	AaA584191 DNA encod
81	72	0.9	4805	5	AaA576462	AaA576462 DNA encod
82	72	0.9	4805	10	ADF59950	ADF59950 Human con
83	72	0.9	4805	10	ADF59987	ADF59987 Human con
84	72	0.9	6507	5	AaA573914	AaA573914 DNA encod
85	72	0.9	6507	5	AaA573607	AaA573607 DNA encod
86	72	0.9	6507	5	AaA592665	AaA592665 DNA encod
87	71	0.9	846	5	AaA571720	AaA571720 DNA encod
88	71	0.9	139573	10	ADH58564	ADH58564 Human Na+
89	71	0.9	156416	13	ABD32817	ABD32817 Human can
90	70	0.9	841	5	AaA576467	AaA576467 DNA encod

# ALIGNMENTS

```

RESULT 1
AAAS9215
ID AAAS9215 standard; DNA; 7582 BP.
AC AAAS9215;
XX
XX 07-NOV-2000 (first entry)
XX Human endogenous retrovirus W (HERV-W) sequence.
XX
XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
XX Human endogenous retrovirus.
XX
XX Key Location/Qualifiers
FH LTR 1..120
FT /*tag= a
FT /note= "R of 5' LTR"
FT LTR 121..575
FT /*tag= b
FT /*tag= "U5 of 5' LTR"
FT primer_bind 579..596
FT /*tag= c
FT CDS 581..7194
FT /*tag= d
FT /note= "ORF1 env538"
FT CDS 7039..7194
FT /*tag= e
FT /note= "ORF2 52 AA"
FT CDS 7112..7255
FT /*tag= f
FT /note= "ORF3 48 AA"
FT misc_feature 7244..7254
FT /*tag= g
FT /note= "polypurine tract"
FT LTR 7256..7582
FT /*tag= h
FT polyA_signal 7563..7569
FT /*tag= i
XX
XX WO200043521-A2.
XX
XX 27-JUL-2000.
XX
XX 21-JAN-2000; 2000WO-FR000144.
XX
XX 21-JAN-1999; 99FR-00000888.
XX (INNR ) BIO MERIEUX.
XX
XX Paranhos-Baccala G, Mallet F, Voisset C;
XX
XX WPI; 2000-499229/44.
XX
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy, contains
PT at least part of the gag gene.
XX
XX Disclosure; Page 49-52; 53pp; French.
XX
XX The present sequence represents an endogenous retrovirus, which is
CC associated with an autoimmune disease, and is integrated into the human
CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
CC proteins derived from it, are useful for diagnosis of autoimmune disease
CC (specifically multiple sclerosis) and for monitoring pregnancy. The
  
```

CC nucleic acid fragments may also be used for in situ labelling of isolated  
 CC chromosomes, while the transcription product can be used to study or  
 CC monitor T cell proliferation in vitro

```

SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;

Query Match 100.0%; Score 7582; DB 3; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAACAATCGGGATATAAAACCCAGGCAATTCGAGCTGGCAACAGAGCCCTTTGGGTCC 60
Db 1 CAACAATCGGGATATAAAACCCAGGCAATTCGAGCTGGCAACAGAGCCCTTTGGGTCC 60

Qy 61 CTTCCTTTGATGGGAGCTGTTTTCATGCTATTTCCTCTTAAATCTTGCACCTGCA 120
Db 61 CTTCCTTTGATGGGAGCTGTTTTCATGCTATTTCCTCTTAAATCTTGCACCTGCA 120

Qy 121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCCTCCACCTGC 180
Db 121 CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCCTCCACCTGC 180

Qy 181 TGTTCACCAACCGCANACCTGCGCTGACTCCCATCCCTCTGATCTCTGAGGCTGC 240
Db 181 TGTTCACCAACCGCANACCTGCGCTGACTCCCATCCCTCTGATCTCTGAGGCTGC 240

Qy 241 CGCTGTGCTCTGATCCAGCGAGCGGCCCATTTGGCGTCCCAATTTGGGCTTAAGGCTTGC 300
Db 241 CGCTGTGCTCTGATCCAGCGAGCGGCCCATTTGGCGTCCCAATTTGGGCTTAAGGCTTGC 300

Qy 301 CATTTGNCCTGACGGCTAGTGCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360
Db 301 CATTTGNCCTGACGGCTAGTGCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360

Qy 361 GGGTTCATGTTCTCTTCTGTACCCACGGCTCTTAATAKAACTATAACACTTACCACA 420
Db 361 GGGTTCATGTTCTCTTCTGTACCCACGGCTCTTAATAKAACTATAACACTTACCACA 420

Qy 421 TGGCCCAAGATTCCATTCCTTTGGAATCGTGAGGSAACGAACTCCAGGTGAGAAATAC 480
Db 421 TGGCCCAAGATTCCATTCCTTTGGAATCGTGAGGSAACGAACTCCAGGTGAGAAATAC 480

Qy 481 GARGCTTGCCACCATCTTTGGAAGCGGCTGCTACCTCTCGGAAGTGGTTACACCACTC 540
Db 481 GARGCTTGCCACCATCTTTGGAAGCGGCTGCTACCTCTCGGAAGTGGTTACACCACTC 540

Qy 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCMRACGGACATCC 600
Db 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCMRACGGACATCC 600

Qy 601 MAAGTGATGGGAAACGTTCCCGCAAGCAAAACGCCCTTAAGACGTATTCTTGGARAAT 660
Db 601 MAAGTGATGGGAAACGTTCCCGCAAGCAAAACGCCCTTAAGACGTATTCTTGGARAAT 660

Qy 661 TGGGACAAATTTGACCTCAGACACTAAGAAGAAACGACTTATATCTTCTGCACTGCC 720
Db 661 TGGGACAAATTTGACCTCAGACACTAAGAAGAAACGACTTATATCTTCTGCACTGCC 720

Qy 721 GCCTGGCAGCTCTGAGGGAAGTATAAATTATAACACCACTTTACAGCTAGACTCTTTTG 780
Db 721 GCCTGGCAGCTCTGAGGGAAGTATAAATTATAACACCACTTTACAGCTAGACTCTTTTG 780

Qy 781 TAGAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAATTAAGACAA 840
Db 781 TAGAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAATTAAGACAA 840

Qy 841 CTCACAAATTATGTAAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 CTCACAAATTATGTAAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 900

Qy 901 CTATCCCAAGCATCCCGGCTCTCTTCCCAATTAAGGACCCCTTTCAACCCAAATGG 960
Db 901 CTATCCCAAGCATCCCGGCTCTCTTCCCAATTAAGGACCCCTTTCAACCCAAATGG 960
  
```

QY 961 TCCAAAAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT 1020  
DB 961 TCCAAAAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT 1020  
QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAAATTCGGCCCGCAGCAGAGTGCATGTGCT 1080  
DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAAATTCGGCCCGCAGCAGAGTGCATGTGCT 1080  
QY 1081 TTTTYYTCTCCAGACTTAAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAAAYCCT 1140  
DB 1081 TTTTYYTCTCCAGACTTAAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAAAYCCT 1140  
QY 1141 GATGGCTATATTTGTGTTTACAGGGTTAGGACAAATTTCTGATCTGACATGGAGAGAT 1200  
DB 1141 GATGGCTATATTTGTGTTTACAGGGTTAGGACAAATTTCTGATCTGACATGGAGAGAT 1200  
QY 1201 ATATATGTCACTGCTAAATCAGACACATAACCCCAATGAGAGAGTGCCACCACTAACTGC 1260  
DB 1201 ATATATGTCACTGCTAAATCAGACACATAACCCCAATGAGAGAGTGCCACCACTAACTGC 1260  
QY 1261 AGCCTGAGRTTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATANGGATGACAACA 1320  
DB 1261 AGCCTGAGRTTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATANGGATGACAACA 1320  
QY 1321 GAAGGAAGANATGATTTCCCAAGCCAGCCAGCAGCGAGTTCCAGTCTASACCTCATTTG 1380  
DB 1321 GAAGGAAGANATGATTTCCCAAGCCAGCCAGCAGCGAGTTCCAGTCTASACCTCATTTG 1380  
QY 1381 GGGACACAGAAATCAGTAACATGGGAGTTGGTGTGTCAGACATTTGCTAACTTGTGTGC 1440  
DB 1381 GGGACACAGAAATCAGTAACATGGGAGTTGGTGTGTCAGACATTTGCTAACTTGTGTGC 1440  
QY 1441 TASAAGGACTTAAGGAAAACCTASGAGAAATCTAYGAAATTTACTCAATGATGTCCACCATA 1500  
DB 1441 TASAAGGACTTAAGGAAAACCTASGAGAAATCTAYGAAATTTACTCAATGATGTCCACCATA 1500  
QY 1501 ACACAGGGNAGGAGAAAATCTTACTGCTTTCTGGAGAGACTAAGGAGGCGATTTGAG 1560  
DB 1501 ACACAGGGNAGGAGAAAATCTTACTGCTTTCTGGAGAGACTAAGGAGGCGATTTGAG 1560  
QY 1561 GAAGCGTGCTCTGTGCTCACTGCTCTGTAAGGCCCACTAATCTTAAAGCGTAAGTT 1620  
DB 1561 GAAGCGTGCTCTGTGCTCACTGCTCTGTAAGGCCCACTAATCTTAAAGCGTAAGTT 1620  
QY 1621 TATCACTCAGTCACTGCTGACATTTAGAAAACCTTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
DB 1621 TATCACTCAGTCACTGCTGACATTTAGAAAACCTTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
QY 1681 CAAAACCTTAGAAAACCTTAGAACTTGGCAACCTTGAAGGAGATCAGGAG 1740  
DB 1681 CAAAACCTTAGAAAACCTTAGAACTTGGCAACCTTGAAGGAGATCAGGAG 1740  
QY 1741 GAGCAGCGGAACAGACAAACGGGATTAAGAAAAGGCCACCGCTTTAGTCAATGACCT 1800  
DB 1741 GAGCAGCGGAACAGACAAACGGGATTAAGAAAAGGCCACCGCTTTAGTCAATGACCT 1800  
QY 1801 CAGGCAAGTGCACTTTGGAGGCTCTGMAAGGGAAGAAAGTGGGCAAAATTTGAATGCTTAA 1860  
DB 1801 CAGGCAAGTGCACTTTGGAGGCTCTGMAAGGGAAGAAAGTGGGCAAAATTTGAATGCTTAA 1860  
QY 1861 TAGGGCTTGCTTCCAGTGGGCTTACAAGGACACTTTAAAAAGATTTGTCAGAGTAGAAG 1920  
DB 1861 TAGGGCTTGCTTCCAGTGGGCTTACAAGGACACTTTAAAAAGATTTGTCAGAGTAGAAG 1920  
QY 1921 TAAGCGCCCTCTGCTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
DB 1921 TAAGCGCCCTCTGCTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
QY 1981 CAGGGACAAAGGCTTTTGGTTCAGAGCCACTAACAGATGATCCAGCAGCAGACTG 2040  
DB 1981 CAGGGACAAAGGCTTTTGGTTCAGAGCCACTAACAGATGATCCAGCAGCAGACTG 2040

QY 2041 AGGGTGCCTGGGCAAGCGGCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTT 2100  
DB 2041 AGGGTGCCTGGGCAAGCGGCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTT 2100  
QY 2101 GACCATTTAGGGCCAGAAAGTTGTCTCTTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
DB 2101 GACCATTTAGGGCCAGAAAGTTGTCTCTTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
QY 2161 CTTTCTGTCCGGCAAACTGTCTCTCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG 2220  
DB 2161 CTTTCTGTCCGGCAAACTGTCTCTCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG 2220  
QY 2221 GGCAGTCACTAGATATCTTTTCCAGCCACTAGTTATGAACTGGGAGCTTTATTTCTT 2280  
DB 2221 GGCAGTCACTAGATATCTTTTCCAGCCACTAGTTATGAACTGGGAGCTTTATTTCTT 2280  
QY 2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTGTTAGGGAGACATTTCTA 2340  
DB 2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTGTTAGGGAGACATTTCTA 2340  
QY 2341 GCAAAAGCAGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTNCC 2400  
DB 2341 GCAAAAGCAGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTNCC 2400  
QY 2401 CCTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGACATATGGAAGAGC 2460  
DB 2401 CCTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGACATATGGAAGAGC 2460  
QY 2461 CAAAGAAATCCCGTCTGTTTCAAGTTAACTAAGGATTTCCACTCTCTTCCCTACCAA 2520  
DB 2461 CAAAGAAATCCCGTCTGTTTCAAGTTAACTAAGGATTTCCACTCTCTTCCCTACCAA 2520  
QY 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
DB 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
QY 2581 AAGCCCAAGGCTTTAGTAAACCAATGCACTCCCTGCGAGTAACTCCGAGTGGATTTGAG 2640  
DB 2581 AAGCCCAAGGCTTTAGTAAACCAATGCACTCCCTGCGAGTAACTCCGAGTGGATTTGAG 2640  
QY 2641 GAGGACACAAAACCCAGTGGAGGAGTGGAGGTTAGTGCAGATCTCAGAGATTTCAATGG 2700  
DB 2641 GAGGACACAAAACCCAGTGGAGGAGTGGAGGTTAGTGCAGATCTCAGAGATTTCAATGG 2700  
QY 2701 AGGCGCTGTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATC 2760  
DB 2701 AGGCGCTGTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATC 2760  
QY 2761 CAGAGGAACAGAGTGGTTTACSTCTGGAACCTTMAAGATGCTTTCTTGTGATCCCTG 2820  
DB 2761 CAGAGGAACAGAGTGGTTTACSTCTGGAACCTTMAAGATGCTTTCTTGTGATCCCTG 2820  
QY 2821 TACATCCTGACTCTCAATTTCTTGTGTTGCTTTGAGATACTTCAAAACCCAGCTCTCAAC 2880  
DB 2821 TACATCCTGACTCTCAATTTCTTGTGTTGCTTTGAGATACTTCAAAACCCAGCTCTCAAC 2880  
QY 2881 TCACCTGGACTTTTACCCCAAGGGTTCCAGGATAGYCCCATCTATTTGGCCAGGAT 2940  
DB 2881 TCACCTGGACTTTTACCCCAAGGGTTCCAGGATAGYCCCATCTATTTGGCCAGGAT 2940  
QY 2941 TAGCCCCAGACTTGAGYCARITYMTATACCTGGAACCTTGTGCTTCTTGTAGTGTGATG 3000  
DB 2941 TAGCCCCAGACTTGAGYCARITYMTATACCTGGAACCTTGTGCTTCTTGTAGTGTGATG 3000  
QY 3001 ATTTACTTTTTRGCGCYRVTTCAGAAAACCTTGTGCACTCAAGCCACCCAGGCTCTTMA 3060  
DB 3001 ATTTACTTTTTRGCGCYRVTTCAGAAAACCTTGTGCACTCAAGCCACCCAGGCTCTTMA 3060  
QY 3061 ATTTCTCTGCACTGTGCTACAGGTTTCCAAACABARCTCARTCTCTGCTCACAGC 3120  
DB 3061 ATTTCTCTGCACTGTGCTACAGGTTTCCAAACABARCTCARTCTCTGCTCACAGC 3120  
QY 3121 AGGTTAAATACCTTAGGCTTAARATTTATCCAAAGGCCACCGGCTCTAGTGAGGAATRYA 3180

|||||  
Db 3121 AGGTTAAATATCTAGGCTTAARATATATCCAAAGGACCCARGCCCTCAGTGAGNAYYA 3180  
|||||  
QY 3181 TCCAGCTATATCTGGCTTATCCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCCTTG 3240  
|||||  
Db 3181 TCCAGCTATATCTGGCTTATCCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCCTTG 3240  
|||||  
QY 3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTTCCCAAGTWTGCCRAATAGCCAGGYCATTA 3300  
|||||  
Db 3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTTCCCAAGTWTGCCRAATAGCCAGGYCATTA 3300  
|||||  
QY 3301 WATACASTAAATTAAGGAACTCAGAAAGCCAAATACCCATTTARTAAAGATGGAAMCTGAA 3360  
|||||  
Db 3301 WATACASTAAATTAAGGAACTCAGAAAGCCAAATACCCATTTARTAAAGATGGAAMCTGAA 3360  
|||||  
QY 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCAGTGTAAAGYT 3420  
|||||  
Db 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCAGTGTAAAGYT 3420  
|||||  
QY 3421 TGCCAAACRGCGCAAGACTTTTSTTATATRTACAGAAAGAAACAGRAAYAGCTCTRGA 3480  
|||||  
Db 3421 TGCCAAACRGCGCAAGACTTTTSTTATATRTACAGAAAGAAACAGRAAYAGCTCTRGA 3480  
|||||  
QY 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAT 3540  
|||||  
Db 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAT 3540  
|||||  
QY 3541 GATGTAGTGGCAAGGGTTGRCYTCAATTTGTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
|||||  
Db 3541 GATGTAGTGGCAAGGGTTGRCYTCAATTTGTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
|||||  
QY 3601 GTATCTGAAGCAGTTAAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
|||||  
Db 3601 GTATCTGAAGCAGTTAAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
|||||  
QY 3661 GTGAAYRGCTACTCACTGCTAAAGAGACTTGTGCTGTACAGACAAACVGTTTACTTAAA 3720  
|||||  
Db 3661 GTGAAYRGCTACTCACTGCTAAAGAGACTTGTGCTGTACAGACAAACVGTTTACTTAAA 3720  
|||||  
QY 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACCTTGTCACCTCTTAAC 3780  
|||||  
Db 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACCTTGTCACCTCTTAAC 3780  
|||||  
QY 3781 CCAGYCNCAATTTCTTCAGACATGAGAAAGATARAAYATACTGTCAACAARTAAAT 3840  
|||||  
Db 3781 CCAGYCNCAATTTCTTCAGACATGAGAAAGATARAAYATACTGTCAACAARTAAAT 3840  
|||||  
QY 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGARGTTCCVTTGACTGATCCYAGCCTT 3900  
|||||  
Db 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGARGTTCCVTTGACTGATCCYAGCCTT 3900  
|||||  
QY 3901 CAACCTTGATATGATGGAAGTTCCTTTGTAGAAAAGACTTTCGAAAAGYGGGGTATGC 3960  
|||||  
Db 3901 CAACCTTGATATGATGGAAGTTCCTTTGTAGAAAAGACTTTCGAAAAGYGGGGTATGC 3960  
|||||  
QY 3961 AGTGTGCTAGTATATGGAATAYTTGAAAGTAAATCCCTCCTCCTCAGGAACTAGTCTYA 4020  
|||||  
Db 3961 AGTGTGCTAGTATATGGAATAYTTGAAAGTAAATCCCTCCTCCTCAGGAACTAGTCTYA 4020  
|||||  
QY 4021 GCTRCAGAACTAATAGCCCTCACTYTKGGCACTAGAAATTTAGGAGAAAGGAAAGGYAAA 4080  
|||||  
Db 4021 GCTRCAGAACTAATAGCCCTCACTYTKGGCACTAGAAATTTAGGAGAAAGGAAAGGYAAA 4080  
|||||  
QY 4081 TATATATACAGACTCTARTATGCTYACCTAGTCTCCTATGCCATGMRGCAATATGSAR 4140  
|||||  
Db 4081 TATATATACAGACTCTARTATGCTYACCTAGTCTCCTATGCCATGMRGCAATATGSAR 4140  
|||||  
QY 4141 AGAAAGGGAATTCCTTAATCTTCGAGRGAAACCTATCATMACATCAGGAAGCCATTAGGAR 4200  
|||||  
Db 4141 AGAAAGGGAATTCCTTAATCTTCGAGRGAAACCTATCATMACATCAGGAAGCCATTAGGAR 4200  
|||||  
QY 4201 ATTATATTATGGCGWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCGYGGGGTCATCA 4260  
|||||

Db 4201 ATTATATTGGCGWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCGYGGGGTCATCA 4260  
|||||  
QY 4261 NAAAGGAAAGBAAGGGAAATASAGRGAAATGSCAAGCAKATATTTGAAGCMAAAGAGC 4320  
|||||  
Db 4261 NAAAGGAAAGBAAGGGAAATASAGRGAAATGSCAAGCAKATATTTGAAGCMAAAGAGC 4320  
|||||  
QY 4321 TGCAAGGCAGACCCCTCCATTAGAAATGCTTTATTAAACTTCCCTTAGTATAGGGTAAATCC 4380  
|||||  
Db 4321 TGCAAGGCAGACCCCTCCATTAGAAATGCTTTATTAAACTTCCCTTAGTATAGGGTAAATCC 4380  
|||||  
QY 4381 CTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGAGG 4440  
|||||  
Db 4381 CTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGAGG 4440  
|||||  
QY 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAGAGAGGGAATACTTTTGCCTGCAAC 4500  
|||||  
Db 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAGAGAGGGAATACTTTTGCCTGCAAC 4500  
|||||  
QY 4501 TATCCAAATGGAAATTTACTTAAACCCCTTCATCAAAACCTTTTCACTTAGGCATCGATAGCAC 4560  
|||||  
Db 4501 TATCCAAATGGAAATTTACTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCAC 4560  
|||||  
QY 4561 CCATCARATGCGCAAAATCATTATTATTACTTGGACAGGCCCTTTTCAAAACCTATCAAGCANAT 4620  
|||||  
Db 4561 CCATCARATGCGCAAAATCATTATTATTACTTGGACAGGCCCTTTTCAAAACCTATCAAGCANAT 4620  
|||||  
QY 4621 AKTCAGGCCCTGTGAATGTGCCARABAAATAATCCCTGCTTATCCCAAGCTCCTTC 4680  
|||||  
Db 4621 AKTCAGGCCCTGTGAATGTGCCARABAAATAATCCCTGCTTATCCCAAGCTCCTTC 4680  
|||||  
QY 4681 AGGARAACAAARAACAGGCCATTCCCTGRARAARACTTGGCAACTGATTTTACCCACAAAG 4740  
|||||  
Db 4681 AGGARAACAAARAACAGGCCATTCCCTGRARAARACTTGGCAACTGATTTTACCCACAAAG 4740  
|||||  
QY 4741 CCCAAACCTCAGGGATTTTCACTAGTCTGGGTARATACCTTTTCAACGGGTTGGGCA 4800  
|||||  
Db 4741 CCCAAACCTCAGGGATTTTCACTAGTCTGGGTARATACCTTTTCAACGGGTTGGGCA 4800  
|||||  
QY 4801 RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCATAGTTTATGNA 4860  
|||||  
Db 4801 RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCATAGTTTATGNA 4860  
|||||  
QY 4861 ATAATCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTCTTTTCCAG 4920  
|||||  
Db 4861 ATAATCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTCTTTTCCAG 4920  
|||||  
QY 4921 GCCACAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
|||||  
Db 4921 GCCACAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
|||||  
QY 4981 TGAAGGCCACAGTCTCAGGGAAGGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
|||||  
Db 4981 TGAAGGCCACAGTCTCAGGGAAGGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
|||||  
QY 5041 AAAAAGCAAAACCCAGGAAACCCACTCACATGCGCTGTCTGTCTGCTATAGCCTTAAAA 5100  
|||||  
Db 5041 AAAAAGCAAAACCCAGGAAACCCACTCACATGCGCTGTCTGTCTGCTATAGCCTTAAAA 5100  
|||||  
QY 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
|||||  
Db 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
|||||  
QY 5161 CCCTTATACCAATGACCTTGTGCTTGAACCAAGACGCAACTTAGTTGACAGATCA 5220  
|||||  
Db 5161 CCCTTATACCAATGACCTTGTGCTTGAACCAAGACGCAACTTAGTTGACAGATCA 5220  
|||||  
QY 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAACAATTACAGAAACCTATCCCTGAGAAG 5280  
|||||  
Db 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAACAATTACAGAAACCTATCCCTGAGAAG 5280  
|||||  
QY 5281 AGGGAAGAAAGAACTATTTCACCCWGTGATGGTATTAGTCAAGTCCCTTCCTCTAAAT 5340  
|||||  
Db 5281 AGGGAAGAAAGAACTATTTCACCCWGTGATGGTATTAGTCAAGTCCCTTCCTCTAAAT 5340  
|||||

QY	5341	CCCATCCCTAGATACATCTCTGGGAGGACCTACCAGTCATTTATYTTACCCCACTG	5400
Db	5341		
QY	5401	CGGTTAAAGTGGCTGGAGTGGATCTTTGGATACATCACCTTGGAGCAAACTCTGGATAC	5460
Db	5401		
QY	5461	TGCCAAGGAACTGGAATCCAGGAGACAAACCTAGCTATTTCTGTGAACCTCTAGAGG	5520
Db	5461		
QY	5521	ATTTGGCCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACATAAATCATAAATCCCC	5580
Db	5521		
QY	5581	ATGSGCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTSTTTCACTCTCACT	5640
Db	5581		
QY	5641	GCACCCCTCCATGCCGCTGTATGACAGTAGCTCCCTYACCMAGAGTTTCTATGGAGA	5700
Db	5641		
QY	5701	ATGCAGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCC	5760
Db	5701		
QY	5761	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCACTCTTTTGCATG	5820
Db	5761		
QY	5821	CATGCAAACTACTCATTTATGGACAGGAAAATGATTAATCTAGTTGTCTTGAGGACTT	5880
Db	5821		
QY	5881	GGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGTCGATGGGGTGAGTT	5940
Db	5881		
QY	5941	CAAGATCAGGCAAGAGAAAAATGTAAGAAAGTAATCTCCCAACTCAACCGGGTACAT	6000
Db	5941		
QY	6001	GGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC	6060
Db	6001		
QY	6061	CATACCTCGCTGGTAAAGCTATTTAATACACCTCACTGGGCTCCATGAGGTCTCGGC	6120
Db	6061		
QY	6121	CAAAACCTTAACTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATC	6180
Db	6121		
QY	6181	CCTGTACCTGAACAATGGAACTTACAGCAGAGAAATAAACACCACTTCCGTTTATGTA	6240
Db	6181		
QY	6241	GGACCTCTGTTTCCAAATSTGGAATTAACCCATACCTCAAACTCACTGTGTAAATTTT	6300
Db	6241		
QY	6301	AGCAATACTACATACACCAACTCCCAATGCATCAGTGGGTAACTCTCCACACAA	6360
Db	6301		
QY	6361	ATAGTCTGCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTTG	6420
Db	6361		

QY	6421	AATGGCTCTTCAGAAATCTATGTGCTTCTCTCTCATTTCTTAGTGCCCCCYATGRCATCTAC	6480
Db	6421		
QY	6481	ACTGAAACAAGATTTATACAGTTATGTCAATCTAAAGCCCGCAACAAAAGAGTACCCATT	6540
Db	6481		
QY	6541	CTTCTCTTTTGTATAGGAGGAGGTGCTAGTGTGCACTAGGTACTGCGGTATC	6600
Db	6541		
QY	6601	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGG	6660
Db	6601		
QY	6661	GTCCGCGACTCCCTGCTCAGCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT	6720
Db	6661		
QY	6721	CRAAATCGAAGAGCTTTTAGACTYGTAAACCGCTGABAGAGGGGAAACCTGTTTATTTT	6780
Db	6721		
QY	6781	GGGGAAGATGCTGTTTATTTATTTATCAATCGGAATCGTCACTGAGAAAGTTAAGAA	6840
Db	6781		
QY	6841	ATTCAGATCGAATCAACGTAKAGCAGARGAGCTTCGAAACACTCGACCTGGGCGCTC	6900
Db	6841		
QY	6901	CTCAGCCRAATGGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG	6960
Db	6901		
QY	6961	CTACTCTCTTTGGACCTGTATCTTTTACCTCTTTTAACTTTCTCTCTCCAGAACT	7020
Db	6961		
QY	7021	GAAGCTGTAAACTCAAAATGAGCCCAAGATCGAGTCCAAAGACTAAGATCTACCCGAGA	7080
Db	7021		
QY	7081	CCCCTGGACCGGCTGYTAGCCACGATCTGATGTTTAAATGAATCAAGACACCCCTCCT	7140
Db	7081		
QY	7141	GAGGAAATCTCAGCTGCAACCTCTATACGCCCCAATTCAGCAGGAAAGCAGTTAGAGC	7200
Db	7141		
QY	7201	GGTGTGGGCAACCTCCCAACAGCAGCTAGGTTTCTGTTGATGGGAGCTGAGA	7260
Db	7201		
QY	7261	GACAGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCYTAAGCCTAGTGTGGAAAGGTG	7320
Db	7261		
QY	7321	ACCAATCAACCTTTAAACAACCGGCTTGCAACTTAGYTACACCTGACCAATCAGAGAG	7380
Db	7321		
QY	7381	CTCACTAAATGCTAATTTAGGCAAGAGAGGAGTTAAGAAATAGCCAATCATYTTATTC	7440
Db	7381		
QY	7441	MTGAGACACACGAGGAGCAATGATCGGGATATAACCCAAAGTTTTCGAGCCGGCAA	7500
Db	7441		
QY	7501	CGGCAACCCCTTTGGTCCCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC	7560

Db 7501 CGGCAACCCCTTGGGTCCTCCCTTGTATGGAGCTCTGTTTTCATGCTATTCAC 7560  
 QY 7561 TCTATTAAATCTGCARCTGCR 7582  
 Db 7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 2  
 AAX25665  
 ID AAX25665 standard; cDNA to mRNA; 7582 BP.  
 AC AAX25665;  
 XX 21-MAY-1999 (first entry)  
 DT Complete human endogenous retrovirus W genome.  
 DE  
 XX Clone; human endogenous retrovirus; genome; autoimmune disease;  
 KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;  
 KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.  
 XX Human endogenous retrovirus.  
 OS  
 XX WO9902696-A1.  
 PN  
 XX 21-JAN-1999.  
 PD  
 XX 06-JUL-1998; 98WO-FR001442.  
 XX 07-JUL-1997; 97FR-00008815.  
 PR  
 XX (INMR ) BIO MERIEUX.  
 PA  
 XX Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;  
 PI WPI; 1999-120897/10.  
 DR  
 XX New nucleic acid sequences from human endogenous retrovirus-W - expressed  
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune  
 PT disease, and abnormal or failed pregnancy.  
 XX  
 XX Claim 1; Page 71-74; 106pp; French.  
 PS  
 CC This sequence represents the complete sequence of the human endogenous  
 CC retrovirus (HERV) W genome. The nucleic acids, their fragments or  
 CC peptides encoded by them are markers of autoimmune disease (e.g. multiple  
 CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,  
 CC insulin-dependent diabetes and related pathologies) and of abnormal or  
 CC unsuccessful pregnancy and can be used as chromosomal markers for  
 CC susceptibility to these conditions, or proximity markers of genes  
 CC associated with this susceptibility  
 XX  
 SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;  
 Query Match 97.3%; Score 7378; DB 2; Length 7582;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7578; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAACATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCCGCCCTTGGGTCC 60  
 Db 1 CAACATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCCGCCCTTGGGTCC 60  
 QY 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTTCACTCTATTAAATCTTCAACTGCA 120  
 Db 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTTCACTCTATTAAATCTTCAACTGCA 120  
 QY 121 CTCCTCTGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCAACCACTGTC 180  
 Db 121 CTCCTCTGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCAACCACTGTC 180  
 QY 181 TGTGTCACCAACCGCANACCTGCGCTGACTCCCATCCCTCTGGATCTCTGAGAGTGC 240

Db 181 TGTGTCACCAACCGCANACCTGCGCTGACTCCATCCCTCTGATCCTGCAAGGTGTC 240  
 QY 241 CGCTGTGCTCTGATCCAGCGARGGCCCATTTGGCGTCCCAATTTGGGCTAAAGCTTGC 300  
 Db 241 CGCTGTGCTCTGATCCAGCGARGGCCCATTTGGCGTCCCAATTTGGGCTAAAGCTTGC 300  
 QY 301 CATTGTNCTCGACGGCTTAAGCTGCTGTTTGTCTAAATTTAGCTGAACACTANTCACT 360  
 Db 301 CATTGTNCTCGACGGCTTAAGCTGCTGTTTGTCTAAATTTAGCTGAACACTANTCACT 360  
 QY 361 GGGTTCCATGTTCTCTCTGTGACCCACGGCTTCTAATAKAACCTATAACACTTACCACA 420  
 Db 361 GGGTTCCATGTTCTCTCTGTGACCCACGGCTTCTAATAKAACCTATAACACTTACCACA 420  
 QY 421 TGGCCCAAGATTCCATTCCTTGGAAATCCGTGAGSCAAAGAACTCCAGGTGAGAAATAC 480  
 Db 421 TGGCCCAAGATTCCATTCCTTGGAAATCCGTGAGSCAAAGAACTCCAGGTGAGAAATAC 480  
 QY 481 GARGCTTCCACCATCTTGGAAAGCGGCTGCTACCTCTTGGAACTGTTTCAACCACTC 540  
 Db 481 GARGCTTCCACCATCTTGGAAAGCGGCTGCTACCTCTTGGAACTGTTTCAACCACTC 540  
 QY 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCAACAGGACATCC 600  
 Db 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCAACAGGACATCC 600  
 QY 601 MAAGTGATGGAAACGTTTCCCGCAAGACAAAAAGCGCCCTTAAGACGTATTCTGGAAAT 660  
 Db 601 MAAGTGATGGAAACGTTTCCCGCAAGACAAAAAGCGCCCTTAAGACGTATTCTGGAAAT 660  
 QY 661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAGAAACGACTTATATTTCTTGCAGTGC 720  
 Db 661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAGAAACGACTTATATTTCTTGCAGTGC 720  
 QY 721 GCCTGGCACTCTGAGGGAAGTATAAAATTATAACACCATCTTTACAGCTAGACTCTTTTG 780  
 Db 721 GCCTGGCACTCTGAGGGAAGTATAAAATTATAACACCATCTTTACAGCTAGACTCTTTTG 780  
 QY 781 TAGAAAAAGGCAATTTGAGTGAAGTGCCTAAGTACAAACTTTTCTTTTCAATTAAGACAA 840  
 Db 781 TAGAAAAAGGCAATTTGAGTGAAGTGCCTAAGTACAAACTTTTCTTTTCAATTAAGACAA 840  
 QY 841 CTCACATTTATGTAATAAAGTCTGATTTATGCTTACAGGAAGCTTTACAGTCTACCTCC 900  
 Db 841 CTCACATTTATGTAATAAAGTCTGATTTATGCTTACAGGAAGCTTTACAGTCTACCTCC 900  
 QY 901 CTATCCAGCATCCCCGACTCTCTTCCCGAMTAAAGGACCCCGCTTCAACCCAAATGG 960  
 Db 901 CTATCCAGCATCCCCGACTCTCTTCCCGAMTAAAGGACCCCGCTTCAACCCAAATGG 960  
 QY 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAATATTTCCCAAT 1020  
 Db 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAATATTTCCCAAT 1020  
 QY 1021 TATGACCCCTCCAGAGCTGGGAGAGAAATTCGGCCCGAGCCAGAGTGCATGTCVT 1080  
 Db 1021 TATGACCCCTCCAGAGCTGGGAGAGAAATTCGGCCCGAGCCAGAGTGCATGTCVT 1080  
 QY 1081 TTTTCTTCCAGACTTAAAGCAATAAAGCAAGACTTGTAGTAAATTTCTCAGATAATCT 1140  
 Db 1081 TTTTCTTCCAGACTTAAAGCAATAAAGCAAGACTTGTAGTAAATTTCTCAGATAATCT 1140  
 QY 1141 GATGCTATATTTGTTTAAAGGGTTAAGGAAATTTTGTGATCTGACATGGAGAT 1200  
 Db 1141 GATGCTATATTTGTTTAAAGGGTTAAGGAAATTTTGTGATCTGACATGGAGAT 1200  
 QY 1201 ATATATGCTACTGCTAAATCAGACTTAACCCCAATGAGAGAGTCCACCACTACCTGC 1260  
 Db 1201 ATATATGCTACTGCTAAATCAGACTTAACCCCAATGAGAGAGTCCACCACTACCTGC 1260  
 QY 1261 AGCTGAGGTTTGGCGATCTCTGCTATCTCAGTCAAGTCAATGGATGATGACAAACA 1320  
 Db 1261 AGCTGAGGTTTGGCGATCTCTGCTATCTCAGTCAAGTCAATGGATGATGACAAACA 1320

QY 1321 GAAGGAAGANAATGATTTCCCAAGCCAGCAGCAGGAGTTCAGTCTASACCTCATTC 1380  
DB 1321 GAAGGAAGANAATGATTTCCCAAGCCAGCAGCAGCAGGAGTTCAGTCTASACCTCATTC 1380  
QY 1381 GGGACACAGAAAATCAGTAACATGGAGATTTGGTCTGCAGACATTTTCTAACTTGTGTGC 1440  
DB 1381 GGGACACAGAAAATCAGTAACATGGAGATTTGGTCTGCAGACATTTTCTAACTTGTGTGC 1440  
QY 1441 TASAAGGACTAAGGAATACTTASGAAGAAATCTAYGAATTTACTCAATGATGTCCACCATA 1500  
DB 1441 TASAAGGACTAAGGAATACTTASGAAGAAATCTAYGAATTTACTCAATGATGTCCACCATA 1500  
QY 1501 ACACAGGGAGGAGAGAAATCCTACTGCTTCTTGGAGAGACTAAGGAGGATTTGAG 1560  
DB 1501 ACACAGGGAGGAGAGAAATCCTACTGCTTCTTGGAGAGACTAAGGAGGATTTGAG 1560  
QY 1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
DB 1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
QY 1621 TATCACTCAGTCACTGAGACATTTAGAAAACTTTCAAAAGTCTGCGTAGGCCCGGAG 1680  
DB 1621 TATCACTCAGTCACTGAGACATTTAGAAAACTTTCAAAAGTCTGCGTAGGCCCGGAG 1680  
QY 1681 CAAAACCTTAGAAACCTTATGAATTTGGCAACTTGGCACTTGGTTTTTTTATATAGAGATCAGGAG 1740  
DB 1681 CAAAACCTTAGAAACCTTATGAATTTGGCAACTTGGCACTTGGTTTTTTTATATAGAGATCAGGAG 1740  
QY 1741 GAGCAGCGGAAACAGGACAAACGGGATTTAAAAAGGCAAAAGCTGGGCAAAATTTGAATGCTTAA 1800  
DB 1741 GAGCAGCGGAAACAGGACAAACGGGATTTAAAAAGGCAAAAGCTGGGCAAAATTTGAATGCTTAA 1800  
QY 1801 CAGGCAAGTGGACTTTTGGAGGCTCTGAAAAGGGAAGGCTGGGCAAAATTTGAATGCTTAA 1860  
DB 1801 CAGGCAAGTGGACTTTTGGAGGCTCTGAAAAGGGAAGGCTGGGCAAAATTTGAATGCTTAA 1860  
QY 1861 TAGGGCTTGTCCAGTGGGTTCTACAGGACACTTTTAAAAAGATTTGCCAAGTAGAAG 1920  
DB 1861 TAGGGCTTGTCCAGTGGGTTCTACAGGACACTTTTAAAAAGATTTGCCAAGTAGAAG 1920  
QY 1921 TAAAGCGCCCTCTGCTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCC 1980  
DB 1921 TAAAGCGCCCTCTGCTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCC 1980  
QY 1981 CAGGGGCAAAAGTCTTTTGAAGTCAAGACCACTAACAGATGATCCAGCAGCAGGACTG 2040  
DB 1981 CAGGGGCAAAAGTCTTTTGAAGTCAAGACCACTAACAGATGATCCAGCAGCAGGACTG 2040  
QY 2041 AGGGTGCCTGGGCAAGCGCAATCCCATGCCATGCCATCACCTCACAGAGCCCTGGGTATGCTT 2100  
DB 2041 AGGGTGCCTGGGCAAGCGCAATCCCATGCCATGCCATCACCTCACAGAGCCCTGGGTATGCTT 2100  
QY 2101 GACCATTTAGGGCCAGAGAGTTGTCTCTGACACTGGTGGGTCTTCTAGTCTTACT 2160  
DB 2101 GACCATTTAGGGCCAGAGAGTTGTCTCTGACACTGGTGGGTCTTCTAGTCTTACT 2160  
QY 2161 CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGAG 2220  
DB 2161 CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCTACTATTTCTGAGGGGTCCNTAAGAG 2220  
QY 2221 GGCAGTCACTAGATACTTTTCCAGCCCACTAAGTTTATGAACCTGGGGAGCTTTATTTCTTT 2280  
DB 2221 GGCAGTCACTAGATACTTTTCCAGCCCACTAAGTTTATGAACCTGGGGAGCTTTATTTCTTT 2280  
QY 2281 TCACATGCTTTCTAATATGCTTTGAAGCCCACTACCTTGTGTAGGGAGAGACATTTCTA 2340  
DB 2281 TCACATGCTTTTCTAATATGCTTTGAAGCCCACTACCTTGTGTAGGGAGAGACATTTCTA 2340  
QY 2341 GCAAAAGCAGGGCCATTTACACCTGAACATAGGAGAGGAAACACCCGTTTGTGTGCC 2400  
DB 2341 GCAAAAGCAGGGCCATTTACACCTGAACATAGGAGAGGAAACACCCGTTTGTGTGCC 2400

QY 2401 CTTGCTTGAGGAAGAAATTAATCCTCAAGTCTGGGCAACAGAAAGCAATATATGACGAGC 2460  
DB 2401 CTTGCTTGAGGAAGAAATTAATCCTCAAGTCTGGGCAACAGAAAGCAATATATGACGAGC 2460  
QY 2461 CAAAGAAATGCCGCTCTGTTCAAGTTTAAACTAAGGATTTCCACTTCTTCCCTACCAA 2520  
DB 2461 CAAAGAAATGCCGCTCTGTTCAAGTTTAAACTAAGGATTTCCACTTCTTCCCTACCAA 2520  
QY 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCAAAGATTTGTTAAGGACTTAA 2580  
DB 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCAAAGATTTGTTAAGGACTTAA 2580  
QY 2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAAGTATTCGTAAGTGTAG 2640  
DB 2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAAGTATTCGTAAGTGTAG 2640  
QY 2641 GAGGCACAGAAACCCAGTGGAGCTGAGAGGTTAGTGCAGATCTCAGGATTTCAATGG 2700  
DB 2641 GAGGCACAGAAACCCAGTGGAGCTGAGAGGTTAGTGCAGATCTCAGGATTTCAATGG 2700  
QY 2701 AGGCCGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATAC 2760  
DB 2701 AGGCCGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATAC 2760  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGAACCTTMMAGGATGCTTCTCTGATCCCTG 2820  
DB 2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGAACCTTMMAGGATGCTTCTCTGATCCCTG 2820  
QY 2821 TACATCCTCAGTCTCAATTTCTTTGTTGCTTTGAAGATCTTCAAAACCCCAATCTCAAC 2880  
DB 2821 TACATCCTCAGTCTCAATTTCTTTGTTGCTTTGAAGATCTTCAAAACCCCAATCTCAAC 2880  
QY 2881 TCACCTGGAATTTTACCCCAAGGTTTCAAGGATAGYCCCATCTATTTTGGCCAGGAT 2940  
DB 2881 TCACCTGGAATTTTACCCCAAGGTTTCAAGGATAGYCCCATCTATTTTGGCCAGGAT 2940  
QY 2941 TAGCCCAAGACTTGAGYCAATYMTCACTCTGACACTCTTGTCTTCTGATGTGATG 3000  
DB 2941 TAGCCCAAGACTTGAGYCAATYMTCACTCTGACACTCTTGTCTTCTGATGTGATG 3000  
QY 3001 ATTTACTTTTTCGCGCTTCTCAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTMA 3060  
DB 3001 ATTTACTTTTTCGCGCTTCTCAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTMA 3060  
QY 3061 ATTTCTCTGACCTGTGGCTACAGGTTTCAAAACSAARGCTCACTCTGCTCAGCAGC 3120  
DB 3061 ATTTCTCTGACCTGTGGCTACAGGTTTCAAAACSAARGCTCACTCTGCTCAGCAGC 3120  
QY 3121 AGGTTAAATACTTAGGCTAABATTTATCAAAGGCCACCAAGGCCCTCAGTGAAGAAAYYA 3180  
DB 3121 AGGTTAAATACTTAGGCTAABATTTATCAAAGGCCACCAAGGCCCTCAGTGAAGAAAYYA 3180  
QY 3181 TCCAGGCTTACTGGCTTATCCTCATCYAAAACCTTAAGCAACTAAGGRRTTCTTCTG 3240  
DB 3181 TCCAGGCTTACTGGCTTATCCTCATCYAAAACCTTAAGCAACTAAGGRRTTCTTCTG 3240  
QY 3241 GORTAAYAGGTTCTGCGAAWATGATTTCCCGAGTWTGGCRAAATACCCAGGYCATTA 3300  
DB 3241 GORTAAYAGGTTCTGCGAAWATGATTTCCCGAGTWTGGCRAAATACCCAGGYCATTA 3300  
QY 3301 WATACASTAATTAAGGAAACTCAGAAAGCCAAATACCCATTTTARTAGATGGAYMCTGAA 3360  
DB 3301 WATACASTAATTAAGGAAACTCAGAAAGCCAAATACCCATTTTARTAGATGGAYMCTGAA 3360  
QY 3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGCCCTTAAACCAAGYCCCAAGTGTTAAGYT 3420  
DB 3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGCCCTTAAACCAAGYCCCAAGTGTTAAGYT 3420  
QY 3421 TCCCAACGGGCAAGACTTTTSTTATATVTCACAGAAAAACAGRAAYAGCTCTRGA 3480  
DB 3421 TCCCAACGGGCAAGACTTTTSTTATATVTCACAGAAAAACAGRAAYAGCTCTRGA 3480  
QY 3481 GTCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCTVTTGGCRYACCTGASTAAGGAAAT 3540



|||||  
3481 GTCCCTACACAGTCTCRAGGGAAGCTTGCCAACTCTGCGCRYACTGASTAAGAAAT 3540  
QY  
3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Db  
3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
QY  
3601 GTATCTGAAGCAGTTTAAATAATATACAGGGRAGAGATCTTACHTGTGTGGACATCTCATGAK 3660  
Db  
3601 GTATCTGAAGCAGTTTAAATAATATACAGGGRAGAGATCTTACHTGTGTGGACATCTCATGAK 3660  
QY  
3661 GTGAAYRGATACACTCACTCTCTAAAGAGACTTGTGCTGTCTAGACAAACVGTCTTAA 3720  
Db  
3661 GTGAAYRGATACACTCACTCTCTAAAGAGACTTGTGCTGTCTAGACAAACVGTCTTAA 3720  
QY  
3721 TTTCAAGGCTCTAATCTTGAARGGCCAGTCTGCRACCTGTGACATCTCTTAAC 3780  
Db  
3721 TTTCAAGGCTCTAATCTTGAARGGCCAGTCTGCRACCTGTGACATCTCTTAAC 3780  
QY  
3781 CCAGYCNCAATTTCTCCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT 3840  
Db  
3781 CCAGYCNCAATTTCTCCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT 3840  
QY  
3841 TCTCAAAACCTATGCCACTCGAGGGACCTTCTGTAGAGTTCCVTTGACTGATCCYAGCCTT 3900  
Db  
3841 TCTCAAAACCTATGCCACTCGAGGGACCTTCTGTAGAGTTCCVTTGACTGATCCYAGCCTT 3900  
QY  
3901 CAACTTGTATCTGATGGAAGTTCCCTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
Db  
3901 CAACTTGTATCTGATGGAAGTTCCCTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
QY  
3961 AGTGGTCAGTGAATAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTGCTYA 4020  
Db  
3961 AGTGGTCAGTGAATAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTGCTYA 4020  
QY  
4021 GCTRCAGACAACTAATAGCCCTCACTGAGGCACTAGAAATAGGAGAGRAAAGGGYAAA 4080  
Db  
4021 GCTRCAGACAACTAATAGCCCTCACTGAGGCACTAGAAATAGGAGAGRAAAGGGYAAA 4080  
QY  
4081 TATATATACAGACTCTTARTATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
Db  
4081 TATATATACAGACTCTTARTATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
QY  
4141 AGAAGGGAATTCCTAACTTCYAGGAAACA CTTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db  
4141 AGAAGGGAATTCCTAACTTCYAGGAAACA CTTATCAMACATCAGGAAGCCATTAGGAR 4200  
QY  
4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGGTGGWAGTCTTACACTGCGYGGGTATCA 4260  
Db  
4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGGTGGWAGTCTTACACTGCGYGGGTATCA 4260  
QY  
4261 NAAAGGAAAGRAAAGGGAATASAGRGAAATGCGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Db  
4261 NAAAGGAAAGRAAAGGGAATASAGRGAAATGCGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
QY  
4321 TGCAAGGCAAGACCTTCCATTTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Db  
4321 TGCAAGGCAAGACCTTCCATTTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
QY  
4381 CTTCCGGGAAACCAAGCCCTAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG 4440  
Db  
4381 CTTCCGGGAAACCAAGCCCTAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG 4440  
QY  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTATTTTGCCTGCAAC 4500  
Db  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTATTTTGCCTGCAAC 4500  
QY  
4501 TATCCAAATGGAATTAATTAACCTTCAATCAACCTTTCACTTTAGGCAATCGATAGCAC 4560  
Db  
4501 TATCCAAATGGAATTAATTAACCTTCAATCAACCTTTCACTTTAGGCAATCGATAGCAC 4560  
QY  
4561 CCATCARATGGCCAAATCATTTTACTCGACAGGCCCTTTTCAAAACTATCAAGCARAT 4620  
|||||

Db  
4561 CCATCARATGGCCAAATCATTTTACTTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
QY  
4621 AKTCAGGCCCTGTGAATGTGCCAPARAAAATAATCCCTGCTCTYATCGCAAGCTCCTTC 4680  
Db  
4621 AKTCAGGCCCTGTGAATGTGCCAPARAAAATAATCCCTGCTCTYATCGCAAGCTCCTTC 4680  
QY  
4681 AGGAAAACAAAARACAGGCCATTAACCTGARAARACCTGGCAACTGATTTTATCCCAACAG 4740  
Db  
4681 AGGAAAACAAAARACAGGCCATTAACCTGARAARACCTGGCAACTGATTTTATCCCAACAG 4740  
QY  
4741 CCCAAACCTCAGGGATTTTCACTATCTACTAGTCTGGGTARATATCTTTTCAACGGTTGGGCA 4800  
Db  
4741 CCCAAACCTCAGGGATTTTCACTATCTACTAGTCTGGGTARATATCTTTTCAACGGTTGGGCA 4800  
QY  
4801 RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Db  
4801 RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
QY  
4861 ATATTTCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Db  
4861 ATATTTCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
QY  
4921 GCCACAGTAAACCCAGGGAGTATCCCAAGGCTTGTAGGTATACGATATCACTTACACTGCGCC 4980  
Db  
4921 GCCACAGTAAACCCAGGGAGTATCCCAAGGCTTGTAGGTATACGATATCACTTACACTGCGCC 4980  
QY  
4981 TGAAGGCCACAGTCTCTCAGGGAAGGTTCAGAAAAATGAATGAAYACTCAAAGGACATCTA 5040  
Db  
4981 TGAAGGCCACAGTCTCTCAGGGAAGGTTCAGAAAAATGAATGAAYACTCAAAGGACATCTA 5040  
QY  
5041 AAAAAGCAAAACCCAGGAAACCCACTCACATGCGCTGCTCTGCTGCTATAGCCTTAAAA 5100  
Db  
5041 AAAAAGCAAAACCCAGGAAACCCACTCACATGCGCTGCTCTGCTGCTATAGCCTTAAAA 5100  
QY  
5101 AGAATCTGCAACTTTTCCCAAAAAGCAGACTTACGCCATACGAAATGCTGTATGGAAGG 5160  
Db  
5101 AGAATCTGCAACTTTTCCCAAAAAGCAGACTTACGCCATACGAAATGCTGTATGGAAGG 5160  
QY  
5161 CCCTTCTAATACCAATGACTTGTGCTTGAACCCCAAGACAGCCAACTTAGTTGCAGACATCA 5220  
Db  
5161 CCCTTCTAATACCAATGACTTGTGCTTGAACCCCAAGACAGCCAACTTAGTTGCAGACATCA 5220  
QY  
5221 CCTCTCTAGCCAAATATCAACAGTCTTTAAACATTAACAGGAACCTTATCCCTGAGAAG 5280  
Db  
5221 CCTCTCTAGCCAAATATCAACAGTCTTTAAACATTAACAGGAACCTTATCCCTGAGAAG 5280  
QY  
5281 AGGAAAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAT 5340  
Db  
5281 AGGAAAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAT 5340  
QY  
5341 CCCCATCTCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATVTATCCCAACTG 5400  
Db  
5341 CCCCATCTCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATVTATCCCAACTG 5400  
QY  
5401 CGGTTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTTGAGTCAATCTTGATAC 5460  
Db  
5401 CGGTTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTTGAGTCAATCTTGATAC 5460  
QY  
5461 TGCCAAAAGGAACCTGAAATCCAGGAGACCAAGCTAGTCTTCTGTGAACTCTTAGAGG 5520  
Db  
5461 TGCCAAAAGGAACCTGAAATCCAGGAGACCAAGCTAGTCTTCTGTGAACTCTTAGAGG 5520  
QY  
5521 ATTTCCGGCTGCTTCTTCAAAACCAACCCAGGAGGAAGTAACATAAATCATAAATCCCCC 5580  
Db  
5521 ATTTCCGGCTGCTTCTTCAAAACCAACCCAGGAGGAAGTAACATAAATCATAAATCCCCC 5580  
QY  
5581 ATGSGCCTCCCTTATCATATTTTCTKTASTGTTSTTTTACCTSTTTCTCTCACT 5640  
Db  
5581 ATGSGCCTCCCTTATCATATTTTCTKTASTGTTSTTTTACCTSTTTCTCTCACT 5640  
QY  
5641 GCACCCCTCCATGCGGCTGTATGACAGTACTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db  
5641 GCACCCCTCCATGCGGCTGTATGACAGTACTCCCTYACCMAGAGTTTCTATGGAGA 5700



XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG04439.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 1; SEQ ID NO 4430; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activities. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic

CC coding sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 7466 BP; 2217 A; 1880 C; 1614 G; 1754 T; 0 U; 1 Other;

Query Match 12.9%; Score 979; DB 5; Length 7466;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1269; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1476 GAATTACTCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC 1535

DB 5407 GAATTACTCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC 5466

QY 1536 TGGAGAGACTAAGGGAGGCAATTGAGGAGCGCTCTCTGTCACCTGACTCTCTGAAG 1595

DB 5467 TGGAGAGACTAAGGGAGGCAATTGAGGAGCGCTCTCTGTCACCTGACTCTCTGAAG 5526

QY 1596 GCCAACTAATCTTAAAGCGTCATGATTTATCACTCAGTCAGTCGACGATTTAGAAAAACT 1655

DB 5527 GCCAACTAATCTTAAAGCGTCATGATTTATCACTCAGTCAGTCGACGATTTAGAAAAACT 5586

QY 1656 TCAAAAGTCTGCGTAGGCGCGAGCAAACTTAGAAACCTTATTGAACCTTGGCACTTC 1715

DB 5587 TCAAAAGTCTGCGTAGGCGCGAGCAAACTTAGAAACCTTATTGAACCTTGGCACTTC 5646

QY 1716 GGTTTTATATAGAGATCAGGAGCAGCGGAGCAACAGGACAACTTAAAAAAA 1775

DB 5647 GGTTTTATATAGAGATCAGGAGCAGCGGAGCAACAGGACAACTTAAAAAAA 5706

QY 1776 AGGCCACCGCTTTAGTCATGACCCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAGGGA 1835

DB 5707 AGGCCACCGCTTTAGTCATGACCCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAGGGA 5766

QY 1836 AAGCTGGGCAAAATTGAATGCTTAATAGGGCTTGCTTCCAGTGGGCTCTACAAGGACACT 1895

DB 5767 AAGCTGGGCAAAATTGAATGCTTAATAGGGCTTGCTTCCAGTGGGCTCTACAAGGACACT 5826

QY 1896 TTAATAAAGATTGTCCCAAGTAGAAGTAGCGCGCCCTTCGTCATGCCCCCTTATTTCAG 1955

DB 5827 TTAATAAAGATTGTCCCAAGTAGAAGTAGCGCGCCCTTCGTCATGCCCCCTTATTTCAG 5886

QY 1956 GGAATCACTGGAAGGCGCCACTGCCCCAGGGGACAAAAGTCTTTTGTAGTCAGAAGCCACTA 2015

DB 5887 GGAATCACTGGAAGGCGCCACTGCCCCAGGGGACAAAAGTCTTTTGTAGTCAGAAGCCACTA 5946

QY 2016 ACCAGATGATCCAGCAGCAGCACTGAGGTCCTGGGGCAAGCGCCATCCCATGCCATCA 2075

DB 5947 ACCAGATGATCCAGCAGCAGCACTGAGGTCCTGGGGCAAGCGCCATCCCATGCCATCA 6006

QY 2076 CCCTCACAGAGCCCTGGGTATGCTTGACCAATTGAGGGCCAGGAAAGTGTCTCTCGGACA 2135

DB 6007 CCCTCACAGAGCCCTGGGTATGCTTGACCAATTGAGGGCCAGGAAAGTGTCTCTCGGACA 6066

QY 2136 CTGCTGGGCTCTCTTAGTCTTACTCTTCTGTCGCGGACAACTGCTCTCCAGATCTGTCA 2195

DB 6067 CTGCTGGGCTCTCTTAGTCTTACTCTTCTGTCGCGGACAACTGCTCTCCAGATCTGTCA 6126

QY 2196 CTATTCTGAGGGGTCCTTAAGAGCGGCGAGTCAGTACTTTTCCAGCCACTTAAGT 2255

DB 6127 CTATTCTGAGGGGTCCTTAAGAGCGGCGAGTCAGTACTTTTCCAGCCACTTAAGT 6185

QY 2256 TATGAACCTGGGGAGCTTTATTCTTTTACATGCTTTTCTAATTAATGCTTGAAGCCCCAC 2315

DB 6186 TATGAACCTGGGGAGCTTTATTCTTTTACATGCTTTTCTAATTAATGCTTGAAGCCCCAC 6245

QY 2316 TACCTTGTAGGGAGAGACATTTAGCAAAAAGAGGGGCGCATTTATACCTTGAAACATAGG 2375

DB 6246 TACCTTGTAGGGAGAGACATTTAGCAAAAAGAGGGGCGCATTTATACCTTGAAACATAGG 6305

QY 2376 AGAGGACACCCGTTGTTGTCGTCGCTTGGAGGAGGAATTAATCCTGAAAGTCTGGG 2435

DB 6306 AGAGGACACCCGTTGTTGTTGT - CCCCTGCTTGAGGAAGGAATTAATCCTGAAAGTCTGGG 6364

QY 2436 CAACAGAGGACAAATATGAGCAGGACCAAGAAATGCCCTCTCTGTTTCAAGTTAAACCTAAG 2495

DB 6365 CAACAGAGGACAAATATGAGCAGGACCAAGAAATGCCCTCTCTGTTTCAAGTTAAACCTAAG 6424

QY 2496 GATTCCACTCTCTTCTTCTTACCAAGGCGAGTACCCCTCAGACCCCAAGGCCCAACAAGA 2555

DB 6425 GATTCCACTCTCTTCTTCTTACCAAGGCGAGTACCCCTCAGACCCCAAGGCCCAACAAGA 6484

QY 2556 TTCCAAAGATTTGTTAGGACTTAAAGCCCAAGGCTTAGTAAACCAATGATGATTAATCC 2615

DB 6485 TTCCAAAGATTTGTTAGGACTTAAAGCCCAAGGCTTAGTAAACCAATGATGATTAATCC 6544

QY 2616 TGCAGTAATTCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGAGCAGTGGAGGGTTAG 2675

DB 6545 TGCAGTAATTCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGAGCAGTGGAGGGTTAG 6604

QY 2676 TGCAGTAATTCAGGATTAATCAATGAGGCGCTTGTCTTTTATATCCAGCTGTACTAGC 2735

DB 6605 TGCAGTAATTCAGGATTAATCAATGAGGCGCTTGTCTTTTATATCCAGCTGTACTAGC 6664

QY 2736 CCTTATCTGTG 2747

DB 6665 CCTTATCTGTG 6676

RESULT 4

AA25664

ID AA25664 standard; cDNA to mRNA; 2372 BP.

XX

AC AA25664;

XX

DT 21-MAY-1999 (first entry)





151 AGGCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTCAAGAA 210  
4862 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGAGCCCTGCTTTCAGG 4921  
211 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGAGCCCTGCTTTCAGG 270  
4922 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 4981  
271 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 330  
4982 GAAGGCCACAGTCTCAGGGAAGGTGAGAAAAATGAATGAAVACTCAAAAGGACATCTAA 5041  
331 GAAGGCCACAGTCTCAGGGAAGGTGAGAAAAATGAATGAAVACTCAAAAGGACATCTAA 390  
5042 AAAAGCAAAACCCAGGAACCCACTACATGCGCTGTCTGTGGCCCTATAGCCCTTAAAAA 5101  
391 AAAAGCAAAACCCAGGAACCCACTACATGCGCTGTCTGTGGCCCTATAGCCCTTAAAAA 450  
5102 GAATCTGCAACTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAGGC 5161  
451 GAATCTGCAACTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAGGC 510  
5162 CTTTCAATAACCAATGACCTTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 5221  
511 CTTTCAATAACCAATGACCTTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 570  
5222 CTCCTTAGCCAAATATCAAGTTCTTAAAAATTAACAAGAACTTATCCCTGGAAGA 5281  
571 CTCCTTAGCCAAATATCAAGTTCTTAAAAATTAACAAGAACTTATCCCTGGAAGA 630  
5282 GCGAAAGAACTATCCACCCWGTGACATGTTAGTCAAGTCCCTTCCTTAATTC 5341  
631 GCGAAAGAACTATCCACCCWGTGACATGTTAGTCAAGTCCCTTCCTTCCTTAATTC 690  
5342 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCAATTTATATACCCCACTGC 5401  
691 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCAATTTATACCCCACTGC 750  
5402 GGTAAAGTGGCTGGAGTGGAGCTTGGATATCATCATCACTTGAGTCAAAATCTCGGATCT 5461  
751 GGTAAAGTGGCTGGAGTGGAGCTTGGATATCATCATCACTTGAGTCAAAATCTCGGATCT 810  
5462 GCCAAGGAACTGAAATCCAGGAGACAGCTAGCTATTCTCTGTAACCTCTAGAGGA 5521  
811 GCCAAGGAACTGAAATCCAGGAGACAGCTAGCTATTCTCTGTAACCTCTAGAGGA 870  
5522 TTTGCGCTCTCTCTCAAAACAAACCCAGGAGAAAGTAACTAAATCATAAATCCCCCA 5581  
871 TTTGCGCTCTCTCTCAAAACAAACCCAGGAGAAAGTAACTAAATCATAAATCCCCCA 929  
5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCCTSTTTTCACTCTCACTG 5641  
930 TGGSCCTCCCTTATCATATTTTCTCTTTACTTGTCTTTTACCCCTCTTTCACTCTCACTG 989  
5642 CACCCCTCCATGCGCTGTATGACAGATAGCTCCCTTACCMAGAGTTTCTATGAGAA 5701  
990 CACCCCTCCATGCGCTGTATGACAGATAGCTCCCTTACCMAGAGTTTCTATGAGAA 1049  
5702 TGCAGGCTCCCGAAATATTGATGCCCATCGATAGGAGTCTTTSTAGGGAAACCCCA 5761  
1050 TGCAGGCTCCCGAAATATTGATGCCCATCGATAGGAGTCTTTCTAAGGGAAACCCCA 1109  
5762 CTTTCACTGCCACACCCCATATGCCCGCACTAGCTATCACTGCGCACTCTTTGATGC 5821  
1110 CTTTCACTGCCACACCCCATATGCCCGCACTAGCTATCACTGCGCACTCTTTGATGC 1169  
5822 ATGCAAAATCTCATTTATGACAGGAAAAATGATTAATCTAGTTGCTCTGGAGGACTTG 5881  
1170 ATGCAAAATCTCATTTATGACAGGAAAAATGATTAATCTAGTTGCTCTGGAGGACTTG 1229  
5882 GAGTCACTGCTGCTGAGCTTACTTACCCAAACTGGTATGCTGTATGGGGTGGAGTTC 5941  
1230 GAGTCACTGCTGCTGAGCTTACTTACCCAAACTGGTATGCTGTATGGGGTGGAGTTC 1289

QY 5942 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCAGGTATCATG 6001  
Db 1290 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCAGGTATCATG 1349  
QY 6002 GCACCTCTAGCCCTTCAAAAGGACTAGTCTCTCAAACTACATGAAACCCCTCGTACCC 6061  
Db 1350 GCACCTCTAGCCCTTCAAAAGGACTAGTCTCTCAAACTACATGAAACCCCTCGTACCC 1409  
QY 6062 ATACTCGCTGTGTAAGCCTATTTAATACCAACCTCTCCTGCGCTCCATGAGGTCTCGGCC 6121  
Db 1410 ATACTCGCTGTGTAAGCCTATTTAATACCAACCTCTCCTGCGCTCCATGAGGTCTCGGCC 1469  
QY 6122 AAAACCCCTACTAATCTGTGTAATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATTC 6181  
Db 1470 AAAACCCCTACTAATCTGTGTAATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATTC 1529  
QY 6182 CTGTACTCTGAACAATGGAACCACTTCAGCACAGAAATTAACACCACTTCCGTTTATAGTAG 6241  
Db 1530 CTGTACTCTGAACAATGGAACCACTTCAGCACAGAAATTAACACCACTTCCGTTTATAGTAG 1589  
QY 6242 GACCTCTGTTTCCAAATSTGGAATTAACCATACCTCAAAACCTCAGCTGTGTAATAATTTA 6301  
Db 1590 GACCTCTGTTTCCAAATCTGGAATTAACCATACCTCAAAACCTCAGCTGTGTAATAATTTA 1649  
QY 6302 GCAATACTACATACACAACTCCCAATGCAATGCAATGAGTGGGTAATCTCTCCACACAAA 6361  
Db 1650 GCAATACTACATACACAACTCCCAATGCAATGCAATGAGTGGGTAATCTCTCCACACAAA 1709  
QY 6362 TAGTCTGCTTACCTCAGGAATATTTTTTGTCTGTGTTACCTCAGGCTATCGTTTGTGTA 6421  
Db 1710 TAGTCTGCTTACCTCAGGAATATTTTTTGTCTGTGTTACCTCAGGCTATCGTTTGTGTA 1769  
QY 6422 ATGGCTCTTCAGAACTATATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6481  
Db 1770 ATGGCTCTTCAGAACTATATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1829  
QY 6482 CTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTC 6541  
Db 1830 CTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTC 1889  
QY 6542 TTCTCTTTTGTATAGAGCAGGAGTCTAGGTGCACTAGGTACTGCGGTATCA 6601  
Db 1890 TTCTCTTTTGTATAGAGCAGGAGTCTAGGTGCACTAGGTACTGCGGTATCA 1949  
QY 6602 CAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
Db 1950 CAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGGG 2009  
QY 6662 TCGCCGACTCCCTGCTCAGCTTTCGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC 6721  
Db 2010 TCGCCGACTCCCTGCTCAGCTTTCGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC 2069

## RESULT 7

ADS30988

ID ADS30988 standard; DNA; 8523 BP.

XX ADS30988;

AC ADS30988;

XX 18-NOV-2004 (first entry)

XX Human genome high complexity repeat found in the HIRA gene #21.

XX Human; ds;

XX histone cell cycle regulation defective, S. cerevisiae homologue A; HIRA;

XX high complexity repeat; in situ hybridisation; Southern blot;

XX chromosome breakpoint; inherited genetic disease; neoplastic disorder;

XX chromosome 22; Digorge syndrome; Velo-Cardio-facial syndrome.

XX Homo sapiens.

XX US2003224356-A1.

PN

XX 04-DEC-2003.  
 XX 14-MAY-2001; 2001US-00854867.  
 XX 16-MAY-2000; 2000US-00573080.  
 XX (KNOLL/) KNOLL J H M.  
 XX (ROGA/) ROGAN P K.  
 XX Knoll JHM, Rogan PK;  
 XX WPI; 2002-062378/08.  
 XX Single copy genomic hybridization probes for detecting specific nucleic  
 XX acid sequences in sample by in situ hybridization useful for detection of  
 XX acquired or inherited genetic diseases.  
 XX Example 1; SEQ ID NO 21; 30pp; English.  
 XX The invention relates to a nucleic acid hybridisation probe comprising a  
 XX labelled, single copy nucleic acid of at least 50 nucleotides, which  
 XX will hybridise to a deduced single copy sequence interval in target  
 XX nucleic acid (TNA) of known sequence. The single copy sequence is deduced  
 XX by comparing the target nucleic acid (e.g. a disease causing gene) with a  
 XX collection of high and low complexity repeat sequences as found in the  
 XX genome of the organism from containing the target nucleic acid. The probe  
 XX is generated by PCR on the target sequence. The probe is essentially free  
 XX of blocking nucleic acid sequences which will hybridise to repeat  
 XX sequences within the genome of which the TNA is a part, and is labelled  
 XX with a label selected from fluorochrome-responsive labels, fluorochromes,  
 XX calorimetric chemical, conjugated proteins, antibodies, antigens and  
 XX their mixtures. The probe is useful in a hybridisation method, where the  
 XX hybridisation method is from in situ hybridisation, Southern blot, and  
 XX other methods in which nucleic acid is immobilised, where the method  
 XX further comprises selecting a single copy nucleic acid which will  
 XX hybridise to a duplication or triplicon sequence domain. The probe is useful  
 XX for determining the existence of previously unknown repeat sequence  
 XX families in a genome. The method comprises reacting a labelled probe with  
 XX the genome, causing the probe to hybridise and ascertaining if the probe  
 XX hybridises to the genome at more than three preferably ten different  
 XX locations as a determination of new repeat sequence family, where the  
 XX determining step comprises selecting the single copy sequence from a  
 XX duplication or triplicon sequence domain. The probe is useful for  
 XX determining a chromosome breakpoint and is useful in the fields for  
 XX cytogenetics and molecular genetics for determining the presence of  
 XX specific nucleic acid sequences in a sample of eukaryotic origin, e.g.  
 XX the probes may be used to analyse specific chromosomal locations by in  
 XX situ hybridisation as a detection of acquired or inherited genetic  
 XX diseases especially for detection of genetic or neoplastic disorders.  
 XX Unlike prior art techniques, the probe permits more precise chromosomal  
 XX breakpoint determinations by in situ hybridisation. The genomic sequence  
 XX comprising the human HIRA gene (histone cell cycle regulation defective,  
 XX S. cerevisiae, homologue A) was analysed for single copy sequence  
 XX intervals for use as probes of the invention. HIRA is located on  
 XX chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and  
 XX Velo-Cardio-facial syndromes. The present sequence is a high complexity  
 XX repeat found within the human genome used to analyse the HIRA gene for  
 XX repeat regions. Note: The sequence data for this patent did not form part  
 XX of the printed specification, but was obtained in electronic format  
 XX directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030224356.  
 XX  
 XX Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;  
 XX  
 XX Query Match 10.3%; Score 779; DB 7; Length 8523;  
 XX Best Local Similarity 98.9%; Pred. No. 0;  
 XX Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 XX  
 XX 4802 AGGCTTCCCTGTAGGACAGAAAAGCCCAAGAGGTAAATAAGGCACTAGTTCATGAAA 4861  
 XX |  
 XX 6071 AGGCTTCCCTGTAGGACAGAAAAGCCCAAGAGGTAAATAAGGCACTAGTTCATGAAA 6130  
 XX |  
 XX 4862 TAATCCAGATTCGGACTTCCCGAGGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 4921

Db	6131	TAATCCAGATTCGGACTTCCCGAGGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG	6190
Qy	4922	CCACAGTAACCCAGGAGTATCCAGCGTTAGGTATAGATATACATTACATCGGCTT	4981
Db	6191	CCACAGTAACCCAGGAGTATCCAGCGTTAGGTATAGATATACATTACATCGGCTT	6250
Qy	4982	GAAGGCCACAGTCTCTCAGGGAAGTTCGAGAAAATGAATGAAYACTCAAGAGACATCTAA	5041
Db	6251	GAAGGCCACAGTCTCTCAGGGAAGTTCGAGAAAATGAATGAAYACTCAAGAGACATCTAA	6310
Qy	5042	AAAAGCAAAACCCAGGAAACCCACCTTCATGCGCTGCTGTTGCTTATAGCTTAAAAA	5101
Db	6311	AAAAGCAAAACCCAGGAAACCCACCTTCATGCGCTGCTGTTGCTTATAGCTTAAAAA	6370
Qy	5102	GAATCTGCAACTTTCCCAAAAAGAGGACTTACGCCATACGAAATGCTGTATGGAAGC	5161
Db	6371	GAATCTGCAACTTTCCCAAAAAGAGGACTTACGCCATACGAAATGCTGTATGGAAGC	6430
Qy	5162	CTTTCATAACCAATGACTTGTGTTGACCCCAAGACAGCCAACTTAGTTGAGACATCAC	5221
Db	6431	CTTTCATAACCAATGACTTGTGTTGACCCCAAGACAGCCAACTTAGTTGAGACATCAC	6490
Qy	5222	CTCCTTAGCCAAATATCAACAAGTTCTTTAAACATTACAAGGAACCTTATCCCTGAGAAG	5281
Db	6491	CTCCTTAGCCAAATATCAACAAGTTCTTTAAACATTACAAGGAACCTTATCCCTGAGAAG	6550
Qy	5282	GGGAAAAGAACTTTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATTC	5341
Db	6551	GGGAAAAGAACTTTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATTC	6610
Qy	5342	CCCATCCCTAGATACATCCTCGGAGGACCTACCCAGTCAATTTTATVATCCCAACTGC	5401
Db	6611	CCCATCCCTAGATACATCCTCGGAGGACCTACCCAGTCAATTTTATVATCCCAACTGC	6670
Qy	5402	GTTTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACACTTGAGTCAAAATCCTGGATCT	5461
Db	6671	GTTTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACACTTGAGTCAAAATCCTGGATCT	6730
Qy	5462	GCCAAAGAACTGAAAATCCAGGAGCAACGCTAGTATTCTCTGTGAACCTCTAGAGA	5521
Db	6731	GCCAAAGAACTGAAAATCCAGGAGCAACGCTAGTATTCTCTGTGAACCTCTCTAGAGA	6790
Qy	5522	TTTGCGCTGCTCTTCAACCAACACACAGGAGGAAGTAACTATAATCATAAATCCCCA	5581
Db	6791	TTTGCGCTGCTCTTCAACCAACACACAGGAGGAAGTAACTATAATCATAAATCCCCA	6849
Qy	5582	TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTTCTCTCTCTCTCTG	5641
Db	6850	TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTTCTCTCTCTCTCTG	6909
Qy	5642	CACCCCTCCATGCGGTATGACAGTAGTCCCTTACCMAGAGTTTCTTATGAGAA	5701
Db	6910	CACCCCTCCATGCGGTATGACAGTAGTCCCTTACCMAGAGTTTCTTATGAGAA	6969
Qy	5702	TGCAGCTCCCGAAATTTGATGCCCATCGTATAGAGTCTTTTSTAGGGAACCCCA	5761
Db	6970	TGCAGCTCCCGAAATTTGATGCCCATCGTATAGAGTCTTTTSTAGGGAACCCCA	7029
Qy	5762	CTTTCACCTGCCACACCCATATGCCCCCACTGCTATCCTCTGCTCTCTCTCTCTCTG	5821
Db	7030	CTTTCACCTGCCACACCCATATGCCCCCACTGCTATCCTCTGCTCTCTCTCTCTG	7089
Qy	5822	ATGCAAAATATCATTTATTTGGACAGAAAATGATTAATCTCTAGTGTCTCTGGAGGACTTG	5881
Db	7090	ATGCAAAATATCATTTATTTGGACAGAAAATGATTAATCTCTAGTGTCTCTGGAGGACTTG	7149
Qy	5882	GAGTCACTGCTGTGGAGTACTTCAACCAACTGGTATGCTCATGGGGTGGAGTTC	5941
Db	7150	GAGTCACTGCTGTGGAGTACTTCAACCAACTGGTATGCTCATGGGGTGGAGTTC	7209
Qy	5942	TAAGATCAGGACAGAGAAAACATGTAAAGAAAGTAAATCTCCCAACTCACCGGGTACATG	6001



```
Db 7210 AAGATCAGGCAAGAGAAAAACATGTAAGAGAGTAATCTCCCACTCACCGGGTACATG 7269
QY 6002 GCACCTCTAGCCCTACAAAGGACTAGATCTTCAAAACTACATGAACCCCTCGTACCC 6061
Db 7270 GCACCTCTAGCCCTACAAAGGACTAGATCTTCAAAACTACATGAACCCCTCGTACCC 7329
QY 6062 ATACTCGCCTGGTAAGCCTATTTAATACCAACCTCTCACTGGGCTCCATGAGGTCTCGGCC 6121
Db 7330 ATACTCGCCTGGTAAGCCTATTTAATACCAACCTCTCACTGGGCTCCATGAGGTCTCGGCC 7389
QY 6122 AAAACCCCTACTAATGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC 6181
Db 7390 AAAACCCCTACTAATGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC 7449
QY 6182 CTGTACTGACCAATGGAACAACTTCAGACAGAAATAAACAACACTTCCTCGTTTGTAGTAG 6241
Db 7450 CTGTACTGACCAATGGAACAACTTCAGACAGAAATAAACAACACTTCCTCGTTTGTAGTAG 7509
QY 6242 GACCTCTGTTTTCCAATSTGGAATAACCATACCTCAAACTCAAACTGCTGTAATAATTTA 6301
Db 7510 GACCTCTGTTTTCCAATSTGGAATAACCATACCTCAAACTCAAACTGCTGTAATAATTTA 7569
QY 6302 GCAATACTACATACACAACTCCCAATGCTAGGTGGGTAACTCTCTCCCAACACAAA 6361
Db 7570 GCAATACTACATACACAACTCCCAATGCTAGGTGGGTAACTCTCTCCCAACACAAA 7629
QY 6362 TAGTCTGCTTACCTCAGGAATATTTTCTGTGTGTAGTCTCAGGCTATCGTGTGTGA 6421
Db 7630 TAGTCTGCTTACCTCAGGAATATTTTCTGTGTGTAGTCTCAGGCTATCGTGTGTGA 7689
QY 6422 ATGCTCTTACAGATCTATGCTCTCTCATCTTCTAGTGGCCCTCATGTCATCTACA 6481
Db 7690 ATGCTCTTACAGATCTATGCTCTCTCATCTTCTAGTGGCCCTCATGTCATCTACA 7749
QY 6482 CTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCTCAAAAGAGTACCCATTC 6541
Db 7750 CTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCTCAAAAGAGTACCCATTC 7809
QY 6542 TTCCTTTTGTATAGGACGAGGAGTGTAGGTGCTAGGTGCTAGTCTGCGGTATCA 6601
Db 7810 TTCCTTTTGTATAGGACGAGGAGTGTAGGTGCTAGGTGCTAGTCTGCGGTATCA 7869
QY 6602 CAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661
Db 7870 CAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 7929
QY 6662 TCGCCGACTCCCTGGTCACTTCGAAAGTCACTTAACTCCCTAGCAGCAGTGTCTTTC 6721
Db 7930 TCGCCGACTCCCTGGTCACTTCGAAAGTCACTTAACTCCCTAGCAGCAGTGTCTTTC 7989

RESULT 8
ADV36376
ID ADY36376 standard; DNA; 8523 BP.
AC ADY36376;
XX
05-MAY-2005 (first entry)
DE HIRA genomic fragment SEQ ID NO 21.
XX
KW hybridization; DNA detection; neoplasm; genetic disorder; cytogenetics;
KW HIRA; ds.
XX
OS Homo sapiens.
XX
FN W0200188089-A2.
XX
PD 22-NOV-2001.
XX
PF 15-MAY-2001; 2001WO-US015674.
XX
PR 16-MAY-2000; 2000US-00573080.
```

PR 14-MAY-2001; 2001US-00854867.

(CHIL-) CHILDREN'S MERCY HOSPITAL.

Knoll JHM, Rogan PK, Cazarro PM;

WPI; 2002-062378/08.

Single copy genomic hybridization probes for detecting specific nucleic acid sequences in sample by in situ hybridization useful for detection of acquired or inherited genetic diseases.

Example 1; SEQ ID NO 21; 67pp; English.

The invention describes a nucleic acid hybridization probe (I) comprising a labeled, single copy nucleic acid of at least 50 nucleotides, which will hybridize to a deduced single copy sequence interval in target nucleic acid (TNA) of known sequence. (I) is useful in a hybridization method which comprises preparing a reaction mixture comprising TNA and (I) which hybridizes to TNA, and causing (I) to hybridize to TNA, where the hybridization method is from in situ hybridization, Southern blot, and other methods in which nucleic acid is immobilized, where the method further comprises selecting a single copy nucleic acid which will hybridize to a duplcon or triplicon sequence domain. (I) is useful for: determining the existence of previously unknown repeat sequence families in a genome; determining a chromosome breakpoint and in the fields of cytogenetics and molecular genetics for determining the presence of specific nucleic acid sequences in a sample of eukaryotic origin, e.g. the probes may be used to analyze specific chromosomal locations by in situ hybridization as a detection of acquired or inherited genetic diseases especially for detection of genetic or neoplastic disorders. Unlike prior art techniques, (I) permits more precise chromosomal breakpoint determinations by in situ hybridization. Hybridization techniques utilizing (I), have made it possible to obtain reliable, easily detectable signals with relatively small probes. A readily detectable signal was obtained with a probe on the order of 2 kb in length, using fluorescent in situ hybridization (FISH) technology. This sensitivity of (I) is improved compared to the prior art, because the probes of (I) are homogeneous single copy sequences. However, smaller amplified segments, each comprising non-repetitive sequences, may also be used in combination as probes to achieve adequate signals for in situ hybridization. Complex single copy probes that hybridize to duplicated or triplicated targets can also increase hybridization signals. This sequence represents a human HIRA genomic sequence that shows homology to a known high-complexity repeat sequence family of the human genome and is used in the creation of an HIRA gene probe.

XX Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;

SQ Query Match 10.3%; Score 779; DB 7; Length 8523;

Best Local Similarity 98.9%; Pred. No. 0;

	Matches 1899;	Conservative	0;	Mismatches	20;	Indels	1;	Gaps	1;
QY	4802	AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAGGCATGATCTATGAAA	4861						
Db	6071	AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAGGCATGATCTATGAAA	6130						
QY	4862	TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTACACATAGCTTCCAGG	4921						
Db	6131	TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTACACATAGCTTCCAGG	6190						
QY	4922	CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCGCT	4981						
Db	6191	CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCGCT	6250						
QY	4982	GAAGGCCACAGTCTCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAGGACATCTAA	5041						
Db	6251	GAAGGCCACAGTCTCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAGGACATCTAA	6310						
QY	5042	AAAAGCAAAACCAGGAACCCACCTCACATGCGCTGTCTGTTCCTATAGCTTAAAAA	5101						
Db	6311	AAAAGCAAAACCAGGAACCCACCTCACATGCGCTGTCTGTTCCTATAGCTTAAAAA	6370						

Qy	5102	GAATCTGGAACCTTTTCCCACAAAAGACAGGACTTAGCCCATACGAAATCTGTTATGGAAGGC	5161
Db	6371	GAATCTGCAACTTTTCCCACAAAAGACAGGACTTAGCCCATACGAAATCTGTTATGGAAGGC	6430
Qy	5162	CCTTCAATACCAATGACTCTTGCTTGCACCCACAGACAGCCAACTTAGTTGTCGACACATCAC	5221
Db	6431	CCTTCAATACCAATGACTCTTGCTTGCACCCACAGACAGCCAACTTAGTTGTCGACACATCAC	6490
Qy	5222	CTCCTTAGCCAAATATCAACAAGTTCCTTAAAAACATTAACAGGAACCTATCCCTGGAAGA	5281
Db	6491	CTCCTTAGCCAAATATCAACAAGTTCCTTAAAAACATTAACAGGAACCTATCCCTGGAAGA	6550
Qy	5282	GGGAAAAGAACTATTCCACCOWGTGACATGTTATAGTCAAGTCCCTTCYCTCTAATTC	5341
Db	6551	GGGAAAAGAACTATTCCACCCTTGTGACATGTTATAGTCAAGTCCCTTCYCTCTAATTC	6610
Qy	5342	CCCATCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATATACCCCAACTGC	5401
Db	6611	CCCATCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATACCCCAACTGC	6670
Qy	5402	GTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATACT	5461
Db	6671	GTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATACT	6730
Qy	5462	GCCAAAGGAACCTGAAATCCAGAGACAACGCTAGCTATTCCTGTGGAACCTCTAGAGGA	5521
Db	6731	GCCAAAGGAACCTGAAATCCAGAGACAACGCTAGCTATTCCTGTGGAACCTCTAGAGGA	6790
Qy	5522	TTTGGGCTGCTCTTCAACAACACAGGAGGAAGTAATAAATCAATAATCCCCCA	5581
Db	6791	TTTGGGCTGCTCTTCAACAACACAGGAGGAAGTAATAAATCAATAATCCCCCA	6849
Qy	5582	TGSGCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCCCTCTTTCACCTCTCACTG	5641
Db	6850	TGSGCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCCTCTTTCACCTCTCACTG	6909
Qy	5642	CACCCCTCCATGCGCTGTATGACAGTAGTCTCCCTCAACMAGAGTTTCTATGGAGAA	5701
Db	6910	CACCCCTCCATGCGCTGTATGACAGTAGTCTCCCTCAACMAGAGTTTCTATGGAGAA	6969
Qy	5702	TGAGAGCTCCGGAATATTTGATGCCCCCATGCTATAGAGTCTTTTAAAGGAACCCCCA	5761
Db	6970	TGAGAGCTCCGGAATATTTGATGCCCCCATGCTATAGAGTCTTTTAAAGGAACCCCCA	7029
Qy	5762	CTTTCACTGCCACACCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGC	5821
Db	7030	CTTTCACTGCCACACCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGC	7089
Qy	5822	ATGCAAACTACTCATTTATTTGGAACAGGAAAAATGATTAATCTCTAGTTGCTCGGAGGACTTG	5881
Db	7090	ATGCAAACTACTCATTTATTTGGAACAGGAAAAATGATTAATCTCTAGTTGCTCGGAGGACTTG	7149
Qy	5882	GAGTCACTGCTGTGTGACTTACTTTCACCCAACTGGTATGCTGTATGGGGGTGGAGTTTC	5941
Db	7150	GAGTCACTGCTGTGTGACTTACTTTCACCCAACTGGTATGCTGTATGGGGGTGGAGTTTC	7209
Qy	5942	AAGATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCACTCACCGSGGTACATG	6001
Db	7210	AAGATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCACTCACCGSGGTACATG	7269
Qy	6002	GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCC	6061
Db	7270	GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCC	7329
Qy	6062	ATACTCGCTGGTAAGCCTATTTAATACACCCCTCACTGGGCTCCATGAGGCTCTCGGCC	6121
Db	7330	ATACTCGCTGGTAAGCCTATTTAATACACCCCTCACTGGGCTCCATGAGGCTCTCGGCC	7389
Qy	6122	AAAACCTACTAACTGTGTGGATAGCCCTCCCTGGAATCTTCARGCCATATGTTTCAATCC	6181
Db	7390	AAAACCTACTAACTGTGTGGATAGCCCTCCCTGGAATCTTCARGCCATATGTTTCAATCC	7449
Qy	6182	CTGTACCTTGAAACAATGGAAACACTTCAGCAACAGAAATAAACACACACTTCGGTTTTAGTAG	6241

Db	7450	CTGTACCTGAACAAATGGAACAACTTCAGCACAGAAATAAAACACCACTTCCTCGTTTATAGTAG	7509
Qy	6242	GACCTCTTGTTTCCAAATSTGGAATAAACCCATACCTCAAAACCTCACCTGTGTAAAAATTTTA	6301
Db	7510	GACCTCTTGTTTCCAAATCTGGAATAAACCCATACCTCAAAACCTCACCTGTGTAAAAATTTTA	7569
Qy	6302	GCAATACTACATACACAAACCAATCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAA	6361
Db	7570	GCAATACTACATACACAAACCAATCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAA	7629
Qy	6362	TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCTATCGTTGTTTGA	6421
Db	7630	TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCTATCGTTGTTTGA	7689
Qy	6422	ATGGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTAGTGCCCCVATGRCATCTTACA	6481
Db	7690	ATGGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTAGTGCCCCCTATGACCATCTACA	7749
Qy	6482	CTGAACAAGATTATACAGTTATGTCTATATCTAAAGCCCGCAACAAAGAAGTACCCATTC	6541
Db	7750	CTGAACAAGATTATACAGTTATGTCTATATCTAAAGCCCGCAACAAAGAAGTACCCATTC	7809
Qy	6542	TTCTTTTGTATAGGACAGGAGTGCTAGGTGCACTAGTACTTGGCATTTGGCGGTATCA	6601
Db	7810	TTCTTTTGTATAGGACAGGAGTGCTAGGTGCACTAGTACTTGGCATTTGGCGGTATCA	7869
Qy	6602	CAACCTCTACTCAGTTCTACTACAAATCTCTCAAGAACTAAATGGGACATGGAACGGG	6661
Db	7870	CAACCTCTACTCAGTTCTACTACAAATCTCTCAAGAACTAAATGGGACATGGAACGGG	7929
Qy	6662	TCGCCGACTCCCTGGTCACTTCGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTC	6721
Db	7930	TCGCCGACTCCCTGGTCACTTCGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTC	7989
RESULT 9			
ID	ABN97929	standard; DNA; 10499 BP.	
XX	ABN97929		
AC	ABN97929;		
DT	01-AUG-2002	(first entry)	
XX	Human retroviral sequence	HERV-7q.	
XX	Autoimmune disease;	HERV-7q; chromosome 7q; immunotherapy;	
KW	multiple sclerosis;	ds.	
XX	Human endogenous retrovirus.		
OS	WO9967395-A1.		
PN	29-DEC-1999.		
PD	23-JUN-1999;	99WO-FR001513.	
XX	23-JUN-1998;	98FR-00007920.	
PR	(INRM ) INSERM INST NAT SANTE & RECH MEDICALE.		
XX	Alliel PM, Perin J, Rieger F;		
PI	WPI; 2000-160587/14.		
DR	New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used		
XX	for diagnosis, treatment and prevention of autoimmune and neurological		
PT	diseases.		
PT	Claim 3; Fig 1; 225pp; French.		
PS	The present invention relates to new nucleic acid sequences of human		
XX	endogenous retrovirus, HERV-7q, which is located on chromosome 7q.		
CC			

CC Regulatory elements associated with HERV-7q may alter expression of other  
 CC genes (even remote genes) on the same chromosome, inducing immunological  
 CC and/or neurological changes (which may be pathological or protective/  
 CC curative). HERV-7q peptides can be used to improve efficiency of the  
 CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding  
 CC sequences can be used in immunogenic or vaccinating compositions, for  
 CC protection against autoimmune diseases, particularly multiple sclerosis.  
 CC The peptides may also be used (by sequence comparison) to detect/identify  
 CC endogenous retroviruses that are abnormally expressed in cancer,  
 CC neuropathologies or other autoimmune diseases. The present sequence was  
 CC used to illustrate the invention  
 CC  
 XX

SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 U; 0 Other;

Query Match 10.3%; Score 779; DB 3; Length 10499;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY	4802	AGGCTTCCCTCTGAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCATGAAA	4861
DB	7101	AGGCTTCCCTCTGAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCATGAAA	7160
QY	4862	TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG	4921
DB	7161	TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG	7220
QY	4922	CCACAGTAACCCAGGAGGTATCCAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG	4981
DB	7221	CCACAGTAACCCAGGAGGTATCCAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG	7280
QY	4982	GAAGGCCACAGTCTCAGGAAAGGTTCGAGAAATGATGAAATCTCAAGGACATCTAA	5041
DB	7281	GAAGGCCACAGTCTCAGGAAAGGTTCGAGAAATGATGAAATCTCAAGGACATCTAA	7340
QY	5042	AAAAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTGCTTATAGCCCTTAAAAA	5101
DB	7341	AAAAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTGCTTATAGCCCTTAAAAA	7400
QY	5102	GAATCTGCAACTTCCCAAAAGAGGAGCTTAGCCCATACGAAATCTGTATGGAAGGC	5161
DB	7401	GAATCTGCAACTTCCCAAAAGAGGAGCTTAGCCCATACGAAATCTGTATGGAAGGC	7460
QY	5162	CCCTTCATACCATGACCTTGCTGTGACCCAGAGCAAGCACTTAGTTGCGACATCAC	5221
DB	7461	CCCTTCATACCATGACCTTGCTGTGACCCAGAGCAAGCACTTAGTTGCGACATCAC	7520
QY	5222	CTCCTTAGCCAAATATCAACAAGTCTTAAACATTTACAGGAACCTATCCCTGAGAGA	5281
DB	7521	CTCCTTAGCCAAATATCAACAAGTCTTAAACATTTACAGGAACCTATCCCTGAGAGA	7580
QY	5282	GGGAAAGAACTATTCCACCMMGTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC	5341
DB	7581	GGGAAAGAACTATTCCACCMMGTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC	7640
QY	5342	CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATYACCCCACTGC	5401
DB	7641	CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATYACCCCACTGC	7700
QY	5402	GTTTAAAGTGGCTGGAGTCTTGGATACATCACCTTGAGTCAAAATCTCGTACT	5461
DB	7701	GTTTAAAGTGGCTGGAGTCTTGGATACATCACCTTGAGTCAAAATCTCGTACT	7760
QY	5462	GCCAAAGGAACCTTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGGA	5521
DB	7761	GCCAAAGGAACCTTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGGA	7820
QY	5522	TTTGGCGCTGCTCTTCAAAACAAACAGGAGGAAAGTAACTAAATCATATAATCCCCCA	5581
DB	7821	TTTGGCGCTGCTCTTCAAAACAAACAGGAGGAAAGTAACTAAATCATATAATCCCCCA	7879
QY	5582	TGSCCTCCCTTACATATTTTCTKTASTGTSTTTTACCTSTTTTCACTCTCACTG	5641
DB	7880	TGSCCTCCCTTACATATTTTCTKTASTGTSTTTTACCTSTTTTCACTCTCACTG	7939

QY	5642	CACCCCTCCATCCCGCTGTATGACACAGTAGTCCCTTACCMAGAGTTTCTTATGGAAA	5701
DB	7940	CACCCCTCCATCCCGCTGTATGACACAGTAGTCCCTTACCMAGAGTTTCTTATGGAAA	7999
QY	5702	TGCAGCGTCCCGAATATTTGATGCCCCCATCTGTATAGGAGTCTTTTAAGGGAAACCCCA	5761
DB	8000	TGCAGCGTCCCGAATATTTGATGCCCCCATCTGTATAGGAGTCTTTTAAGGGAAACCCCA	8059
QY	5762	CTTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGC	5821
DB	8060	CTTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGC	8119
QY	5822	ATGCAATACTCAATTTGGACAGGAAAAATGATTAATCTTAGTTGTCTGGAGGACTTG	5881
DB	8120	ATGCAATACTCAATTTGGACAGGAAAAATGATTAATCTTAGTTGTCTGGAGGACTTG	8179
QY	5882	GAGTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGTATGCGGGTGGAGTTC	5941
DB	8180	GAGTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGTATGCGGGTGGAGTTC	8239
QY	5942	AAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCGGGTATCATG	6001
DB	8240	AAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCGGGTATCATG	8299
QY	6002	GCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTTACATGAACCCCTCCGATCCC	6061
DB	8300	GCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTTACATGAACCCCTCCGATCCC	8359
QY	6062	ATACTCGCTTGTGTAAGCTTATTAATACACCCCTCACTGGGCTTCATAGAGTCTCGGCC	6121
DB	8360	ATACTCGCTTGTGTAAGCTTATTAATACACCCCTCACTGGGCTTCATAGAGTCTCGGCC	8419
QY	6122	AAAAACCTACTACTTCTGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC	6181
DB	8420	AAAAACCTACTACTTCTGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC	8479
QY	6182	CTGTACTGAACTGAAACCAACTTTCAGCACAGAAATAAACCACTTCGGTTTATAGTAG	6241
DB	8480	CTGTACTGAACTGAAACCAACTTTCAGCACAGAAATAAACCACTTCGGTTTATAGTAG	8539
QY	6242	GACCTCTGTTTCCAAATSTGGAATAACCATACCTCAAACTCACCTGTGTAAAAATTTA	6301
DB	8540	GACCTCTGTTTCCAAATSTGGAATAACCATACCTCAAACTCACCTGTGTAAAAATTTA	8599
QY	6302	GCAATACTACATACACCAACTCCCAATGATCAGTGGGTAACTCCCTCCACACAAA	6361
DB	8600	GCAATACTACATACACCAACTCCCAATGATCAGTGGGTAACTCCCTCCACACAAA	8659
QY	6362	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTGA	6421
DB	8660	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTGA	8719
QY	6422	ATGGCTCTTTCAGAACTATGCTTCTCTCAATTTAGTGCCCTCCATATGRCATCTACA	6481
DB	8720	ATGGCTCTTTCAGAACTATGCTTCTCTCAATTTAGTGCCCTCCATATGRCATCTACA	8779
QY	6482	CTGAAACAGATTTATACAGTTATGATATCTAAGCCCGCAACAAAGAGTACCCATTC	6541
DB	8780	CTGAAACAGATTTATACAGTTATGATATCTAAGCCCGCAACAAAGAGTACCCATTC	8839
QY	6542	TTCCCTTTTGTATAGGACAGGAGTCTAGGTGCACTAGGTACTGTCATTTGGCGGTATCA	6601
DB	8840	TTCCCTTTTGTATAGGACAGGAGTCTAGGTGCACTAGGTACTGTCATTTGGCGGTATCA	8899
QY	6602	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG	6661
DB	8900	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG	8959
QY	6662	TGCGCGACTCCCTGGTCCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTC	6721
DB	8960	TGCGCGACTCCCTGGTCCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTC	9019

RESULT 10  
 ABL61744  
 ID ABL61744 standard; DNA; 56093 BP.  
 XX AC ABL61744;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Colon adenocarcinoma related gene sequence SEQ ID NO: 81.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US010838.  
 XX  
 PR 05-JUN-2000; 2000US-0209473P.  
 PR 05-JUN-2000; 2000US-0209531P.  
 PR 18-SEP-2000; 2000US-0233133P.  
 PR 18-SEP-2000; 2000US-0233617P.  
 PR 20-SEP-2000; 2000US-0234009P.  
 PR 20-SEP-2000; 2000US-0234034P.  
 PR 20-SEP-2000; 2000US-0234052P.  
 PR 22-SEP-2000; 2000US-0234509P.  
 PR 22-SEP-2000; 2000US-0234567P.  
 PR 25-SEP-2000; 2000US-0234923P.  
 PR 25-SEP-2000; 2000US-0234924P.  
 PR 25-SEP-2000; 2000US-0235077P.  
 PR 25-SEP-2000; 2000US-0235082P.  
 PR 25-SEP-2000; 2000US-0235134P.  
 PR 26-SEP-2000; 2000US-0235280P.  
 PR 26-SEP-2000; 2000US-0235637P.  
 PR 26-SEP-2000; 2000US-0235638P.  
 PR 27-SEP-2000; 2000US-0235711P.  
 PR 27-SEP-2000; 2000US-0235720P.  
 PR 27-SEP-2000; 2000US-0235840P.  
 PR 27-SEP-2000; 2000US-0235863P.  
 PR 28-SEP-2000; 2000US-0236028P.  
 PR 28-SEP-2000; 2000US-0236032P.  
 PR 28-SEP-2000; 2000US-0236033P.  
 PR 28-SEP-2000; 2000US-0236034P.  
 PR 28-SEP-2000; 2000US-0236109P.  
 PR 28-SEP-2000; 2000US-0236111P.  
 PR 29-SEP-2000; 2000US-0236842P.  
 PR 29-SEP-2000; 2000US-0236891P.  
 PR 02-OCT-2000; 2000US-0237172P.  
 PR 02-OCT-2000; 2000US-0237173P.  
 PR 02-OCT-2000; 2000US-0237278P.  
 PR 02-OCT-2000; 2000US-0237294P.  
 PR 02-OCT-2000; 2000US-0237295P.  
 PR 02-OCT-2000; 2000US-0237316P.  
 PR 03-OCT-2000; 2000US-0237425P.  
 PR 03-OCT-2000; 2000US-0237598P.  
 PR 03-OCT-2000; 2000US-0237604P.  
 PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX WPI; 2002-188264/24.  
 XX

XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX  
 PS Claim 1; SEQ ID NO 81; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
 CC tumour  
 XX  
 SQ Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;  
 Query Match 10.3%; Score 779; DB 6; Length 56093;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 QY 4802 AGGCTTCCCTGTAGGACAGAAAGGCCCGAGGTAAATGAAGGCACTAGTTCATGAAA 4861  
 DB 35101 AGGCTTCCCTGTAGGACAGAAAGGCCCGAGGTAAATGAAGGCACTAGTTCATGAAA 35160  
 QY 4862 TAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 4921  
 DB 35161 TAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 35220  
 QY 4922 CCACAGTAACCCAGGAGTATCCCAAGGCTTAGGTATACGATATCACTTACATCGGCCT 4981  
 DB 35221 CCACAGTAACCCAGGAGTATCCCAAGGCTTAGGTATACGATATCACTTACATCGGCCT 35280  
 QY 4982 GAAGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAA 5041  
 DB 35281 GAAGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAA 35340  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGCTGTGTCTATAGCCTTAAAAA 5101  
 DB 35341 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGCTGTGTCTATAGCCTTAAAAA 35400  
 QY 5102 GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161  
 DB 35401 GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 35460  
 QY 5162 CCTTCATTAACCAATGACCTTGCTTGACCAAGCAGCACTTAGTTGCGAGACATCAC 5221  
 DB 35461 CCTTCATTAACCAATGACCTTGCTTGACCAAGCAGCACTTAGTTGCGAGACATCAC 35520  
 QY 5222 CTCCTTAGCCAAATATCAACAGTTCTTAAAAACATTTACAAGGAACCTATCCCTGAGAAGA 5281  
 DB 35521 CTCCTTAGCCAAATATCAACAGTTCTTAAAAACATTTACAAGGAACCTATCCCTGAGAAGA 35580  
 QY 5282 GGGAAAAGAACTATTCCACCCWGTGACATGTATAGTCAAGTCCCTTCCTTAATTC 5341  
 DB 35581 GGGAAAAGAACTATTCCACCCCTTGTGACATGTATAGTCAAGTCCCTTCCTTAATTC 35640  
 QY 5342 CCCATCCCTAGATACATCTCTGGGAGGACCTTACCAGTCATTTATYACCCCACTGC 5401  
 DB 35641 CCCATCCCTAGATACATCTCTGGGAGGACCTTACCAGTCATTTATYACCCCACTGC 35700  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATCTGAGTCAAACTCTCGGATACT 5461

```
Db 35701 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCCTGGATACT 35760
Qy 5462 GCCAAGGAACCTGAAAATCCAGAGACAAGCTAGCTATTCCTGTGAAACCTCTAGAGGA 5521
Db 35761 GCCAAGGAACCTGAAAATCCAGAGACAAGCTAGCTATTCCTGTGAAACCTCTAGAGGA 35820
Qy 5522 TTTCGGCTCTCTTTCAACAACAACACAGAGAGAAAGTAATAAATCAATAAATCCCCCA 5581
Db 35821 TTTCGGCTCTCTTTCAACAACAACACAGAGAGAAAGTAATAAATCAATAAATCCCCCA 35879
Qy 5582 TGSCCTCCCTTATCATATTTTCTCTKTAATGTSTTTTACCTCTTTTCACTCTCACTG 5641
Db 35880 TGSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCTCTTTTCACTCTCACTG 35939
Qy 5642 CACCCCTCCATGCCCTGTATGACCAAGTAGTCCCTCACTGAGTCTTTTAAAGGAACCCCA 5701
Db 35940 CACCCCTCCATGCCCTGTATGACCAAGTAGTCCCTCACTGAGTCTTTTAAAGGAACCCCA 35999
Qy 5702 TGCAGGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAACCCCA 5761
Db 36000 TGCAGGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAACCCCA 36059
Qy 5762 CTTTCACTGCCACACCCATATGATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 5821
Db 36060 CTTTCACTGCCACACCCATATGATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 36119
Qy 5822 ATGCAATACTCTATTATGACAGAGAAATGATTAATCCTAGTTGCTCGGAGGACTTG 5881
Db 36120 ATGCAATACTCTATTATGACAGAGAAATGATTAATCCTAGTTGCTCGGAGGACTTG 36179
Qy 5882 GAGTCACTGCTCTGTGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTTC 5941
Db 36180 GAGTCACTGCTCTGTGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTTC 36239
Qy 5942 AAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCCAACTCACCGGGGTACATG 6001
Db 36240 AAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCCAACTCACCGGGGTACATG 36299
Qy 6002 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC 6061
Db 36300 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC 36359
Qy 6062 ATACTCGCTGGTAAGCCTATTTAATACCACTCTACTGGGCTCCATGAGGTCTCGGCC 6121
Db 36360 ATACTCGCTGGTAAGCCTATTTAATACCACTCTACTGGGCTCCATGAGGTCTCGGCC 36419
Qy 6122 AAAACCTCTACTACTGTGTGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 6181
Db 36420 AAAACCTCTACTACTGTGTGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 36479
Qy 6182 CTGTACTGAACAAATGGAACTTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAG 6241
Db 36480 CTGTACTGAACAAATGGAACTTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAG 36539
Qy 6242 GACCTCTGTTTCCATSTGGAATTAACCATACCTCAAACTCACTGCTGTGTAATAATTA 6301
Db 36540 GACCTCTGTTTCCATSTGGAATTAACCATACCTCAAACTCACTGCTGTGTAATAATTA 36599
Qy 6302 GCAATACTACATACACCAACTCCCAATGCATCAGTGGTAACTCTCCACACAAA 6361
Db 36600 GCAATACTACATACACCAACTCCCAATGCATCAGTGGTAACTCTCCACACAAA 36659
Qy 6362 TAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTGA 6421
Db 36660 TAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTGTACCTCAGCCTATCGTTGTTGA 36719
Qy 6422 ATGGCTCTTCAGAAATCTATGTGCTCTCTCAFTTCTAGTGCCCTCCATGCCATCTACA 6481
Db 36720 ATGGCTCTTCAGAAATCTATGTGCTCTCTCTCAFTTCTAGTGCCCTCCATGCCATCTACA 36779
Qy 6482 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGGCAACAAAGAGTACCATTC 6541
Db 36780 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGGCAACAAAGAGTACCATTC 36839
```

---

```
Qy 6542 TTCTTTTGTATTAGAGCAGGAGTCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATCA 6601
Db 36840 TTCTTTTGTATTAGAGCAGGAGTCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATCA 36899
Qy 6602 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6661
Db 36900 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGG 36959
Qy 6662 TGCCCGACTCCCTGGTCACTTTCGAAGATCAACTTAACTCCCTAGACAGTAGTCTCTTC 6721
Db 36960 TGCCCGACTCCCTGGTCACTTTCGAAGATCAACTTAACTCCCTAGACAGTAGTCTCTTC 37019

RESULT 11
AAS84210
ID AAS84210 standard; cDNA; 6394 BP.
XX
AC AAS84210;
XX
DT 13-FEB-2002 (first entry)
XX
DNA encoding novel human diagnostic protein #20014.
XX
Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
(HYSE-) HYSEQ INC.
XX
Drmanac RT, Liu C, Tang YT;
XX
WPI; 2001-639362/73.
DR P-PSDB; ABG20023.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.
XX
Claim 1; SEQ ID NO 20014; 103pp; English.
PS
The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (II) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful for treating disorders
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have application of mutations
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
coding sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
CC
```

XX	SQ	Sequence 6394 BP; 1840 A; 1597 C; 1384 G; 1571 T; 0 U; 2 Other;	
		Query Match 9.9%; Score 750; DB 5; Length 6394;	
		Best Local Similarity 99.8%; Pred. No. 0;	
		Matches 920; Conservative 1; Mismatches 1; Indels 1; Gaps 1;	
QY	1476	GAATTACTCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC	1535
DB	3463	GAATTACTCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC	3522
QY	1536	TGGAGAGACTAAGGGAGGCAATGAGGAGCGTCTCTCTGTCACCTGACTTCTTGAAG	1595
DB	3523	TGGAGAGACTAAGGGAGGCAATGAGGAGCGTCTCTCTGTCACCTGACTTCTTGAAG	3582
QY	1596	GCCAACTAATCTTAAAGCTTAATGTTATCATCTCAGTCAGCTGCAGACATTAGAAAAACT	1655
DB	3583	GCCAACTAATCTTAAAGCTTAATGTTATCATCTCAGTCAGCTGCAGACATTAGAAAAACT	3642
QY	1656	TCAAAAGTCTGCGTAGGCGCGGAGCAAACTTAGAAACCCCTATTGAACTTGGCAACTC	1715
DB	3643	TCAAAAGTCTGCGTAGGCGCGGAGCAAACTTAGAAACCCCTATTGAACTTGGCAACTC	3702
QY	1716	GGTTTTTATATAGAGATCAGGAGCAGCGGAAACAGGACAAACGGGATTAAAAAAA	1775
DB	3703	GGTTTTTATATAGAGATCAGGAGCAGCGGAAACAGGACAAACGGGATTAAAAAAA	3762
QY	1776	AGGCCAACCGCTTTAGTCATGACCCCTCAGGCAAGTGGACTTTGGAGGCTCTGAAAAGGGA	1835
DB	3763	AGGCCAACCGCTTTAGTCATGACCCCTCAGGCAAGTGGACTTTGGAGGCTCTGAAAAGGGA	3822
QY	1836	AAAGCTGGGCAATTTGAATGCCCTTAATAGGCTTGTCTTCAGTCCGCTCTACAGGACACT	1895
DB	3823	AAAGCTGGGCAATTTGAATGCCCTTAATAGGCTTGTCTTCAGTCCGCTCTACAGGACACT	3882
QY	1896	TTAAAAAGATTGTCCAAAGTAGAAGTAAGCCGCCCTTCGTCCATGCCCTTTATTTCAAG	1955
DB	3883	TTAAAAAGATTGTCCAAAGTAGAAGTAAGCCGCCCTTCGTCCATGCCCTTTATTTCAAG	3942
QY	1956	GGAATCACTGGAAGGCCACTGCCCCAGGGGACAAAGGTCCTTTTGTAGTCAGAGGCACATA	2015
DB	3943	GGAATCACTGGAAGGCCACTGCCCCAGGGGACAAAGGTCCTTTTGTAGTCAGAGGCACATA	4002
QY	2016	ACCAGATGATCAGCAGCAGGACTGAGGTGCTTGGGGCAAGCGCCATCCATGCCATCA	2075
DB	4003	ACCAGATGATCAGCAGCAGGACTGAGGTGCTTGGGGCAAGCGCCATCCATGCCATCA	4062
QY	2076	CCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAGTTGCTCTCTGGACA	2135
DB	4063	CCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAGTTGCTCTCTGGACA	4122
QY	2136	CTGGTGGGTCTTCTTGTAGTCTTCTTGTCTCCGGGCAACTGTCCTCAGATCTGTCA	2195
DB	4123	CTGGTGGGTCTTCTTGTAGTCTTCTTGTCTCCGGGCAACTGTCCTCAGATCTGTCA	4182
QY	2196	CTATTCTGAGGGGTCCTNTAAGACGGGAGTCACTAGATACCTTTTCCAGGCACCTAAGT	2255
DB	4183	CTATTCTGAGGGGTCCTNTAAGACGGGAGTCACTAGATACCTTTTCCAGGCACCTAAGT	4241
QY	2256	TATGAATCGGGAGCTTTATCTTTTTCATATGCTTTTCAATTTGCTTGAAGCCCAAC	2315
DB	4242	TATGAATCGGGAGCTTTATCTTTTTCATATGCTTTTCAATTTGCTTGAAGCCCAAC	4301
QY	2316	TACCTTTGTTAGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATACCTGAACATAGG	2375
DB	4302	TACCTTTGTTAGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATACCTGAACATAGG	4361
QY	2376	AGAAGGAACACCCGTTTGTGT 2397	
DB	4362	AGAAGGAACACCCGTTTGTGT 4383	

AED10984	
ID	AED10984 standard; DNA; 1614 BP.
XX	AC AED10984;
XX	15-DEC-2005 (first entry)
XX	HERV R393E/Q envelope protein DNA SEQ ID NO 115.
DE	virucide; cytostatic; immunosuppressive; antiallergic; vaccine;
XX	screening; immune modulation; graft rejection; allergy;
KW	autoimmune disease; cancer; viral infection; envelope; env; gene; ds;
KW	mutant.
XX	Human endogenous retrovirus W.
XX	WO2005095442-A1.
XX	13-OCT-2005.
XX	30-MAR-2005; 2005WO-EP0033339.
XX	30-MAR-2004; 2004EP-00290838.
XX	(INSTR ) INST ROUSSY GUSTAVE.
PA	(CNRS ) CENT NAT RECH SCI.
PA	(UTPA-) UNIV PARIS SUD.
XX	Renard M, Mangeney M, Heidmann T;
XX	WPI; 2005-703311/72.
DR	P-P5DB; AED10985.
XX	New polypeptide capable of modulating the immunosuppressive properties of a viral protein or its fragment, useful for the manufacture of a vaccine for treating or preventing viral infections or cancer.
PT	Claim 41; SEQ ID NO 115; 262pp; English.
XX	The invention describes a polypeptide of 7-20 amino acid residues, which is capable of modulating the immunosuppressive properties of a viral protein or its fragment, against the host in which it is expressed (immunosuppression-modulatory sequence) when it substitutes the homologous sequence of the viral protein or its fragment, the polypeptide comprising the minimum following consensus amino acid sequence of (I).
CC	Also described are: a mutated ENV protein resulting from the mutation of a wild type ENV protein essentially carrying the following sequence: NY 1 Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic acid coding for the polypeptide or the mutated ENV protein; an eukaryotic or prokaryotic expression vector comprising the nucleic acid as well as the elements necessary for the expression of the nucleic acid; a recombinant cell comprising the nucleic acid or the eukaryotic or prokaryotic expression vector; a pharmaceutical or a vaccine composition comprising as active substance; at least one polypeptide, mutated ENV protein or its fragments, nucleic acid, prokaryotic or eukaryotic expression vector, or recombinant cell; in association with a pharmaceutically acceptable carrier; and antibodies or their fragments, scfv polypeptides, aptamers, or binding peptides, directed against mutated ENV proteins, provided that the antibodies or their fragments, scfv polypeptides, aptamers, or binding peptides do not bind to the corresponding wild type ENV proteins. The first mutation of a first amino acid and optionally of a second mutation of a second amino acid in a wild type viral envelope (ENV) protein is useful for manufacturing a mutated ENV protein having a modified immunosuppressive activity with respect to the wild type ENV protein. The protein is useful for the manufacture of a medicament or a vaccine intended for the prevention and/or the treatment of viral diseases, such as HTLV or FeLV infections; cancer; pathologies requiring an inhibition of the immune system, such as autoimmune diseases, allergies or graft rejections. The protein is also useful for screening compounds liable to modulate the immunosuppressive activity of viruses or tumor cells, where the compounds to screen are peptides, in particular peptides comprising from 5-30 amino acids, such as peptides originating from combinatorial peptide libraries. This sequence encodes

CC HERV envelope protein.

Sequence 1614 BP; 439 A; 452 C; 294 G; 426 T; 0 U; 3 Other;

Query Match 9.6%; Score 731; DB 14; Length 1614;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0

5685	QY	AGATGTTCTTATGGAGAAATGCAGCGTCCCGGAAATATTATGATGCCCATCTGATAGAGTCT	5744
105	Db	AGATGTTCTTATGGAGAAATGCAGCGTCCCGGAAATATTATGATGCCCATCTGATAGAGTCT	164
5745	QY	TTTAAAGGGAAACCCCACTTCACTGTCGCCACACCCATATGCCCGCGAACTGCTATCACTC	5804
165	Db	TTCTAAGGAAACCCCACTTCACTGTCGCCACACCCATATGCCCGCGAACTGCTATCACTC	224
5805	QY	TGCCACTCTTTTGCATGTCACAAATACCTCATTTATGGACAGGAAAAATGATTAATCCTAG	5864
225	Db	TGCCACTCTTTTGCATGTCACAAATACCTCATTTATGGACAGGAAAAATGATTAATCCTAG	284
5865	QY	TTTGCTCGAGGACTTGGAGTCACTGCTGTGTGGACTTACTTCCACCAAACTGGTATGTC	5924
285	Db	TTTGCTCGAGGACTTGGAGTCACTGCTGTGTGGACTTACTTCCACCAAACTGGTATGTC	344
5925	QY	TGATGGGGTGGAGTTCCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	5984
345	Db	TGATGGGGTGGAGTTCCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	404
5985	QY	ACTCACCGGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA	6044
405	Db	ACTCACCGGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA	464
6045	QY	TGAAACCTTCGTACCCATFACTGCTGGTGAAGCTATTTAATACACCTCTCTGGGCT	6104
465	Db	TGAAACCTTCGTACCCATFACTGCTGGTGAAGCTATTTAATACACCTCTCTGGGCT	524
6105	QY	CCATGAGGTCCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCCCTGAACCTTCAR	6164
525	Db	CCATGAGGTCCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCCCTGAACCTTCAG	584
6165	QY	GCCATATGTTTCAATCCCTGTACTGAAACAATGGAAACACTTCAGCACAGAAATAAACAC	6224
585	Db	GCCATATGTTTCAATCCCTGTACTGAAACAATGGAAACACTTCAGCACAGAAATAAACAC	644
6225	QY	CACCTCGGTTTTAGTAGGACCTCTTGTTTCCAAATGGAATAAACCCATACCTCAAACT	6284
645	Db	CACCTCGGTTTTAGTAGGACCTCTTGTTTCCAAATGGAATAAACCCATACCTCAAACT	704
6285	QY	CACCTGTGAAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGT	6344
705	Db	CACCTGTGAAAAATTTAGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGT	764
6345	QY	AACTCTCCCAACAATAAGTCTGCCPAACCTCAGGAATAATTTTGTCTGTGGTACCTC	6404
765	Db	AACTCTCCCAACAATAAGTCTGCCPAACCTCAGGAATAATTTTGTCTGTGGTACCTC	824
6405	QY	AGCCTATCGTGTGTTGAAATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTAGTGCC	6464
825	Db	AGCCTATCGTGTGTTGAAATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTAGTGCC	884
6465	QY	CCCYATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAA	6524
885	Db	CCCTATGACCATCTACACTGAAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAA	944
6525	QY	CAAAAGATACCCATTCTTCCTTTTGTTATAGGACAGGAGTGTAGGTGCACTAGGTAC	6584
945	Db	CAAAAGATACCCATTCTTCCTTTTGTTATAGGACAGGAGTGTAGGTGCACTAGGTAC	1004
6585	QY	TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTACTCTCAAGAACTTAA	6644
1005	Db	TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTACTCTCAAGAACTTAA	1064
6645	QY	TGGGGAATGGAAACGGGTCCGCACTCCCTGGTCACTCTGCAAGATCAACTTAACCTCCCT	6704

Db  
1065 TGGGGCATGSAACGGGTCCGACTCCCTGGTGACCTTGCAGAGATCAACTTAATCCCTCT  
Qy  
6705 AGCAGCAGTAGTCCTTC 6721  
Db  
1125 AGCAGCAGTAGTCCTTC 1141

RESULT 13  
AED10986

AED10986  
ID AED10986 standard; DNA; 1614 BP.

XX

AC AED10986;

5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528

DT 15-DEC-2005 (first entry)  
vv

DE  
XX  
HERV R393E/Q F399A envelope protein DNA SEQ ID NO 117.

XX  
KW virucide; cytostatic; immunosuppressive; antiallergic; vaccine;  
KW screening; immune modulation; graft rejection; allergy;  
KW autoimmune disease; cancer; viral infection; envelope; env; gen  
KW mutant.  
KW

XX Human endogenous retrovirus W. OS

XX  
PN WO2005095442-A1

PD 13-OCT-2005.

30-MAR-2005; 2005WO-EP0033339.

XX  
PR 30-MAR-2004; 2004EP-00290838.

XX  
PA (INSR ) INST ROUSSY GUSTAVE.

PA (CNRS ) CENT NAT RECH SCI.  
PA (UTPA-) UNIV PARIS SUD.

XX  
PI Renard M, Mangeney M, Heidmann T;

XX  
DR WPI; 2005-703311/72.

DR	P-PSDB; AED10987.	
XX		
XX		
PT	New polypeptide capable of modulating the immunosuppressive properties of a viral protein or its fragment, useful for the manufacture of a vaccine for treating or preventing viral infections or cancer.	
PT		
PT		

XX  
PS Claim 41; SEQ ID NO 117; 262bp; English.

The invention describes a polypeptide of 7-20 amino acid residues, which is capable of modulating the immunosuppressive properties of a viral protein or its fragment, against the host in which it is expressed (immunosuppression-modulatory sequence) when it substitutes the homologous sequence of the viral protein or its fragment, the polypeptide comprising the minimum following consensus amino acid sequence of (I). Also described are: a mutated ENV protein resulting from the mutation of a wild type ENV protein essentially carrying the following sequence: NY 1 Y 2 Y 3 L Y 4 Y 5 L Y 6 Y 7 Y 8 X Y 9 Y 10 Y 11 C Y 12 X 2 (II); a nucleic acid coding for the polypeptide or the mutated ENV protein; an eukaryotic or prokaryotic expression vector comprising the nucleic acid as well as the elements necessary for the expression of the nucleic acid; a recombinant cell comprising the nucleic acid or the eukaryotic or prokaryotic expression vector; a pharmaceutical or a vaccine composition comprising as active substance: at least one polypeptide, mutated ENV protein or its fragments, nucleic acid, prokaryotic or eukaryotic expression vector, or recombinant cell; in association with a pharmaceutically acceptable carrier; and antibodies or their fragments, scFv polypeptides, aptamers, or binding peptides, directed against mutated ENV proteins, provided that the antibodies or their fragments, scFv polypeptides, aptamers, or binding peptides do not bind to the corresponding wild type ENV proteins. The first mutation of a first amino acid and optionally of a second mutation of a second amino acid in a wild type viral envelope (ENV) protein is useful for manufacturing a mutated ENV protein having a modified immunosuppressive activity with respect to



CC the wild type ENV protein. The protein is useful for the manufacture of a  
CC medicament or a vaccine intended for the prevention and/or the treatment  
CC of viral diseases, such as HTLV or FeLV infections; cancer; pathologies  
CC requiring an inhibition of the immune system, such as autoimmune  
CC diseases, allergies or graft rejections. The protein is also useful for  
CC screening compounds liable to modulate the immunosuppressive activity of  
CC viruses or tumor cells, where the compounds to screen are peptides, in  
CC particular peptides comprising from 5-30 amino acids, such as peptides  
CC originating from combinatorial peptide libraries. This sequence encodes  
CC HERV envelope protein.  
XX

SQ Sequence 1614 BP; 439 A; 452 C; 294 G; 423 T; 0 U; 6 Other;

Query Match 9.6%; Score 731; DB 14; Length 1614;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	5685	AGAGTTTCTATGGAGAAATGACGCTCCGCGAAATATTGATGCCCATCGTATAGGAGTCT	5744
Db	105	AGAGTTTCTATGGAGAAATGACGCTCCGCGAAATATTGATGCCCATCGTATAGGAGTCT	164
QY	5745	TTSTAAGGGAACCCCACTTCACTGCGCCACACCATATGCCCGCAACTGCTATCACTC	5804
Db	165	TTSTAAGGGAACCCCACTTCACTGCGCCACACCATATGCCCGCAACTGCTATCACTC	224
QY	5805	TGCCCACTCTTTCATGCGATGCAAAATACCTATTATTGGACAGAAATAATGATTATCTCTAG	5864
Db	225	TGCCCACTCTTTCATGCGATGCAAAATACCTATTATTGGACAGAAATAATGATTATCTCTAG	284
QY	5865	TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTC	5924
Db	285	TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTC	344
QY	5925	TGATGGGGTGGAGTTCAAGATCAGCGAAGAGAAAAACATGTAAGAAGTAATCTCCCA	5984
Db	345	TGATGGGGTGGAGTTCAAGATCAGCGAAGAGAAAAACATGTAAGAAGTAATCTCCCA	404
QY	5985	ACTCACCGGGTACATGGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA	6044
Db	405	ACTCACCGGGTACATGGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA	464
QY	6045	TGAACCCCTCCGTACCCATCTCGCTCGTGAAGCCCTATTATATACACCCCTCACTGGGCT	6104
Db	465	TGAACCCCTCCGTACCCATCTCGCTCGTGAAGCCCTATTATATACACCCCTCACTGGGCT	524
QY	6105	CCATGAGGTCCTCGGCCCAAAACCTACTAACTGTGGATATGCCCTCCCGCTGAACTTCAR	6164
Db	525	CCATGAGGTCCTCGGCCCAAAACCTACTAACTGTGGATATGCCCTCCCGCTGAACTTCAG	584
QY	6165	GCCATATGTTTCAATCCCTGTACTTGAACAAATGGAACAACTTCAGCAGAAATAAACAC	6224
Db	585	GCCATATGTTTCAATCCCTGTACTTGAACAAATGGAACAACTTCAGCAGAAATAAACAC	644
QY	6225	CACCTTCGGTTTATAGGAGCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACT	6284
Db	645	CACCTTCGGTTTATAGGAGCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACT	704
QY	6285	CACCTGTGTAAATTTAGCAATFACATACATACACCAACCTCCCAATGCATCAGTGGGT	6344
Db	705	CACCTGTGTAAATTTAGCAATFACATACATACACCAACCTCCCAATGCATCAGTGGGT	764
QY	6345	AACTCTCCACACAAATAGTCTGCTACCCCTCAGGAATATTTTGTGTGTGTACCTC	6404
Db	765	AACTCTCCACACAAATAGTCTGCTACCCCTCAGGAATATTTTGTGTGTGTACCTC	824
QY	6405	AGCCTATCGTTCTTTGAATGGCTCTTTCAGAACTATGTCTCTCAATCTTATGTGCC	6464
Db	825	AGCCTATCGTTCTTTGAATGGCTCTTTCAGAACTATGTCTCTCAATCTTATGTGCC	884
QY	6465	CCCATATGCACTTCACTGAAACAGATTTATACAGTTATGTATATCTTAAGCCCCGCAA	6524
Db	885	CCCATATGCACTTCACTGAAACAGATTTATACAGTTATGTATATCTTAAGCCCCGCAA	944

QY	6525	CAAAAGAGTACCCATTTCTTCTTTTATAGGACGAGGTGCTAGGTGCACCTAGGTAC	6584
Db	945	CAAAAGAGTACCCATTTCTTCTTTTATAGGACGAGGTGCTAGGTGCACCTAGGTAC	1004
QY	6585	TGGCATTTGGCGGTATACAACTCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA	6644
Db	1005	TGGCATTTGGCGGTATACAACTCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA	1064
QY	6645	TGGGAGCATGGAACGGGTGCGCGACTCCCTGGTCACCTTGCAGATCAACTTAACCTCCCT	6704
Db	1065	TGGGAGCATGGAACGGGTGCGCGACTCCCTGGTCACCTTGCAGATCAACTTAACCTCCCT	1124
QY	6705	AGCAGCAGTAGTCTCTTC 6721	
Db	1125	AGCAGCAGTAGTCTCTTC 1141	

RESULT 14

AAH20070

ID AAH20070 standard; DNA; 1617 BP.

XX AAH20070;

AC AAH20070;

DT 08-AUG-2001 (first entry)

XX HERV-W envelope protein G coding sequence.

DE Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;  
KW envelope protein; multiple sclerosis-related superantigen; vaccine;  
KW surface antigen; transmembrane; multiple sclerosis; neuroprotective;  
KW antisense-therapy; autoimmune disorder; ds.

OS Human endogenous retrovirus.

XX Key Location/Qualifiers

FT CDS 1..1617

FT /tag= a

FT /product= "HERV-W envelope protein G"

XX WO200131021-A1.

XX 03-MAY-2001.

XX 30-OCT-2000; 2000WO-EP010659.

XX 28-OCT-1999; 99EP-00402690.

XX (UYGE-) UNIV GENEVE.

XX Conrad B, Mach B;

XX WPI; 2001-316336/33.

XX P-PSDB; AAB75138.

XX New human retrovirus HERV-W ENV proteins/peptides having superantigen  
XX activity useful for diagnosing and treating multiple sclerosis.

XX Claim 13; Fig 10; 94pp; English.

XX On the basis of the PBS t-RNA motif used for the classification of human  
XX endogenous retrovirus (HERVs) the full length endogenous provirus which  
XX was located on the long arm of human chromosome 7 (7q21-22) has been  
XX designated HERV-W. The present invention describes proteins or peptides  
XX (I) having superantigen (SAG) activity comprising the ENV protein (ENV)  
XX of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)  
XX have neuroprotective activity, and can be used in vaccines; antisense-  
XX therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are  
XX useful for diagnosing multiple sclerosis (MS) or HERV-W-associated  
XX disorders. (I) are also useful for identifying substances (and optionally  
XX recovering) capable of binding to a retroviral superantigen associated  
XX with MS, substances capable of blocking SAG activity and substances  
XX capable of blocking transcription or translation of HERV-W retroviral  
XX superantigen. A protein or peptide derived from (I), modified to be

CC devoid of SAg activity and being capable of generating an immune response  
CC against HERV-W retroviral SAg is useful in therapy. Nucleic acid  
CC molecules encoding (I) are useful as vaccines against MS. Substances  
CC capable of blocking SAg activity, capable of binding to a retroviral  
CC superantigen associated with MS, or capable of blocking transcription or  
CC translation of HERV-W retroviral superantigen for use in treating or  
CC preventing MS, obtained using (I) are useful for the treatment and  
CC prevention of MS. (I) and nucleic acids encoding them are useful for  
CC diagnosing autoimmune disease. The present sequence encodes the  
CC specifically claimed envelope protein of HERV-W designated G  
XX  
SQ Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 U; 0 Other;

Query Match 9.68; Score 731; DB 5; Length 1617;  
Best Local Similarity 99.48; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGATGCGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
DB |||||  
105 AGAGTTTCTATGGAGATGCGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 164  
QY 5745 TTSTAAGGGAACCCCACTTCTACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804  
DB |||||  
165 TTCTAAGGGAACCCCACTTCTACTGCCACACCCATATGCCCGCAACTGCTATCACTC 224  
QY 5805 TGGCACTCTTTGATGATGATCAATATCTATTATGGACAGGAAATGATTATCTCTAG 5864  
DB |||||  
225 TGGCACTCTTTGATGATGATCAATATCTATTATGGACAGGAAATGATTATCTCTAG 284  
QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGCTGTGTGGACTTACTTCCACCAAACTGTTATGTC 5924  
DB |||||  
285 TTGTCTCTGGAGGACTTGGAGTCACTGCTGTGTGGACTTACTTCCACCAAACTGTTATGTC 344  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAATATCTCCCA 5984  
DB |||||  
345 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAATATCTCCCA 404  
QY 5985 ACTCACGGGTGATAGGACCTCTAGCCCTTCAAGAGCTAGATCTCTCAAAACTACA 6044  
DB |||||  
405 ACTCACGGGTGATAGGACCTCTAGCCCTTCAAGAGCTAGATCTCTCTCAAAACTACA 464  
QY 6045 TGAACCTCTCGTACCATCTCGCTTGGTAAGCTATTATTAACCACTCTACTGGCT 6104  
DB |||||  
465 TGAACCTCTCGTACCATCTCGCTTGGTAAGCTATTATTAACCACTCTACTGGCT 524  
QY 6105 CCATGAGGTCTCGGCCCAAAACCTCTACTGTTGGATATGCTCCCTGAACTTCAR 6164  
DB |||||  
525 CCATGAGGTCTCGGCCCAAAACCTCTACTGTTGGATATGCTCCCTGAACTTCAG 584  
QY 6165 GCCATATGTTCAATCGCTGTACTGAAATGGAACAACTTCAGCAGCAAGAAATAACAC 6224  
DB |||||  
585 GCCATATGTTCAATCGCTGTACTGAAATGGAACAACTTCAGCAGCAAGAAATAACAC 644  
QY 6225 CACTTCGGTTTATAGTAGGACCTCTGTTTCCATSTGGAATACCCATACCTCAACCT 6284  
DB |||||  
645 CACTTCGGTTTATAGTAGGACCTCTGTTTCCATCTGGAATACCCATACCTCAACCT 704  
QY 6285 CACTGTGTAAATTTAGCAATACTATACACAAACCACTCCCAATGATCAGTGGGT 6344  
DB |||||  
705 CACTGTGTAAATTTAGCAATACTATACACAAACCACTCCCAATGATCAGTGGGT 764  
QY 6345 AACTCTCCCAACAAATAGTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 6404  
DB |||||  
765 AACTCTCCCAACAAATAGTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 824  
QY 6405 AGCCTATGTTGTTGATGGCTCTTCAAGATCTATGCTCTCTCTCATCTTAGTGCC 6464  
DB |||||  
825 AGCCTATGTTGTTGATGGCTCTTCAAGATCTATGCTCTCTCTCATCTTAGTGCC 884  
QY 6465 CCCATGRCCTCTACATGAAACAGATTTATACAGTTATGTCATATCTAAGCCCGCAA 6524  
DB |||||  
885 CCCATGRCCTCTACATGAAACAGATTTATACAGTTATGTCATATCTAAGCCCGCAA 944

QY 6525 CAAAAGAGTACCATTCTTCTTTTGTATTAGGAGCAGGAGTCTAGGTGCTAGGTAC 6584  
DB |||||  
945 CAAAAGAGTACCATTCTTCTTTTGTATTAGGAGCAGGAGTCTAGGTGCTAGGTAC 1004  
QY 6585 TGGCATTTGGCGGTATCACAACTCTACTAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
DB |||||  
1005 TGGCATTTGGCGGTATCACAACTCTACTAGTTCTACTACAACTATCTCAAGAACTAAA 1064  
QY 6645 TGGGGACATGGAACGGGTGCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCT 6704  
DB |||||  
1065 TGGGGACATGGAACGGGTGCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCT 1124  
QY 6705 AGCAGCAGTAGTCTTTC 6721  
DB |||||  
1125 AGCAGCAGTAGTCTTTC 1141

## RESULT 15

AED10942

ID AED10942 standard; DNA; 1617 BP.

XX AED10942;

XX AC AC  
XX XX  
DT 15-DEC-2005 (first entry)

XX HERV envelope protein DNA SEQ ID NO 73.

XX viricide; cytostatic; immunosuppressive; antiallergic; vaccine;  
KW screening; immune modulation; graft rejection; allergy;  
KW autoimmune disease; cancer; viral infection; envelope; env; gene; ds.

XX Human endogenous retrovirus W.

XX WO2005095442-A1.

XX 13-OCT-2005.

XX 30-MAR-2005; 2005WO-BP003339.

XX 30-MAR-2004; 2004EP-00290838.

XX (INSR ) INST ROUSSY GUSTAVE.  
PA (CNRS ) CENT NAT RECH SCI.  
PA (UYPA-) UNIV PARIS SUD.

XX Renard M, Mangeney M, Heidmann T;

XX WPI; 2005-703311/72.

XX P-PSDB; AED10943.

XX New polypeptide capable of modulating the immunosuppressive properties of  
PT a viral protein or its fragment, useful for the manufacture of a vaccine  
PT for treating or preventing viral infections or cancer.

XX Disclosure; SEQ ID NO 73; 262pp; English.

XX The invention describes a polypeptide of 7-20 amino acid residues, which  
CC is capable of modulating the immunosuppressive properties of a viral  
CC protein or its fragment, against the host in which it is expressed  
CC (immunosuppression-modulatory sequence) when it substitutes the  
CC homologous sequence of the viral protein or its fragment, the polypeptide  
CC comprising the minimum following consensus amino acid sequence of (I).  
CC Also described are: a mutated ENV protein resulting from the mutation of (I).  
CC a wild type ENV protein essentially carrying the following sequence: NY 1  
CC Y 2 Y 3 IY 4 Y 5 IY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic  
CC acid coding for the polypeptide or the mutated ENV protein; an eukaryotic  
CC or prokaryotic expression vector comprising the nucleic acid as well as  
CC the elements necessary for the expression of the nucleic acid; a  
CC recombinant cell comprising the nucleic acid or the eukaryotic or  
CC prokaryotic expression vector; a pharmaceutical or a vaccine composition  
CC comprising as active substance: at least one polypeptide, mutated ENV  
CC protein or its fragments, nucleic acid, prokaryotic or eukaryotic  
CC expression vector, or recombinant cell; in association with a

CC pharmaceutically acceptable carrier; and antibodies or their fragments,  
 CC scFv polypeptides, aptamers, or binding peptides, directed against  
 CC mutated ENV proteins, provided that the antibodies or their fragments,  
 CC scFv polypeptides, aptamers, or binding peptides do not bind to the  
 CC corresponding wild type ENV proteins. The first mutation of a first amino  
 CC acid and optionally of a second mutation of a second amino acid in a wild  
 CC type viral envelope (ENV) protein is useful for manufacturing a mutated  
 CC ENV protein having a modified immunosuppressive activity with respect to  
 CC the wild type ENV protein. The protein is useful for the manufacture of a  
 CC medicament or a vaccine intended for the prevention and/or the treatment  
 CC of viral diseases, such as HTLV or Fely infections; cancer; pathologies  
 CC requiring an inhibition of the immune system, such as autoimmune  
 CC diseases, allergies or graft rejections. The protein is also useful for  
 CC screening compounds liable to modulate the immunosuppressive activity of  
 CC viruses or tumor cells, where the compounds to screen are peptides, in  
 CC particular peptides comprising from 5-30 amino acids, such as peptides  
 CC originating from combinatorial peptide libraries. This sequence encodes  
 CC HERV envelope protein.

XX Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 U; 0 Other;

Query Match 9.6%; Score 731; DB 14; Length 1617;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGATGACGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
 Db 105 AGAGTTTCTATGGAGATGACGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 164  
 QY 5745 TTSTAAGGGAACCCACCTTCACTGCGCACACCCATATGCCCCCACTGCTACACTC 5804  
 Db 165 TTCTAAGGGAACCCACCTTCACTGCGCACACCCATATGCCCCCACTGCTACACTC 224  
 QY 5805 TGCCACTCTTTGATGATGCAATTAATCTATTATTGGACAGGAAATGATTAATCCTAG 5864  
 Db 225 TGCCACTCTTTGATGATGCAATTAATCTATTATTGGACAGGAAATGATTAATCCTAG 284  
 QY 5865 TTGCTCTGGAGACTGGAGTCACTGTCGTGGGACTTACTTCCACCAACTGGTATGTC 5924  
 Db 285 TTGCTCTGGAGACTGGAGTCACTGTCGTGGGACTTACTTCCACCAACTGGTATGTC 344  
 QY 5925 TCATGGGGTGAGTTCAAGATCAGCAGAGAGAGAAACATGTAAGAAGTATCTCCCA 5984  
 Db 345 TCATGGGGTGAGTTCAAGATCAGCAGAGAGAGAAACATGTAAGAAGTATCTCCCA 404  
 QY 5985 ACTACCCGGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 6044  
 Db 405 ACTACCCGGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 464  
 QY 6045 TGAACCCCTCCGTACCCATCTAGCGCTGTGAAGCCTATTATATACACCCCTCACTGGGT 6104  
 Db 465 TGAACCCCTCCGTACCCATCTAGCGCTGTGAAGCCTATTATATACACCCCTCACTGGGT 524  
 QY 6105 CCATGAGTCTCGGCCCAAAACCTACTTAATCTGTTGGATATGCTCCCTCGAATTCAR 6164  
 Db 525 CCATGAGTCTCGGCCCAAAACCTACTTAATCTGTTGGATATGCTCCCTCGAATTCAG 584  
 QY 6165 GCATATGTTTCAATCCCTGTACCTGAACAATGGAAACAATTTTTCAGCACAGAAATAAACAC 6224  
 Db 585 GCATATGTTTCAATCCCTGTACCTGAACAATGGAAACAATTTTTCAGCACAGAAATAAACAC 644  
 QY 6225 CACTTCCGTTTTAGTAGGACCTTTGTTTCCAATSTGGAATAAACCCATACCTCAAACT 6284  
 Db 645 CACTTCCGTTTTAGTAGGACCTTTGTTTCCAATSTGGAATAAACCCATACCTCAAACT 704  
 QY 6285 CACTTGTGTAATTTAGCAATACTACATACACCAACTCCCAATGATCAGTGGGT 6344  
 Db 705 CACTTGTGTAATTTAGCAATACTACATACCAACTCCCAATGATCAGTGGGT 764  
 QY 6345 AACTCTCCACCAAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTTC 6404  
 Db 765 AACTCTCCACCAAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTTC 824

QY 6405 AGCCTATCGTTGTTTGAATGGCTCTTCAAGAACTATATGCTCTCTCTCATTTAGTGCC 6464  
 Db 825 AGCCTATCGTTGTTTGAATGGCTCTTCAAGAACTATATGCTCTCTCTCATTTAGTGCC 884  
 QY 6465 CCCATGRCCTATACACTGAAACAGATTTATACAGTTTATGTCATATCTAAGCCCCGCAA 6524  
 Db 885 CCCTATGACCATCTACACTGAAACAGATTTATACAGTTTATGTCATATCTAAGCCCCGCAA 944  
 QY 6525 CAAAAGAGTACCCATCTTCTCTTTTATAGGAGCAGAGTGTAGTGACCTAGGTAC 6584  
 Db 945 CAAAAGAGTACCCATCTTCTCTTTTATAGGAGCAGAGTGTAGTGACCTAGGTAC 1004  
 QY 6585 TGCGATTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
 Db 1005 TGGCATTTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1064  
 QY 6645 TGGGACATGAAAGGGGTGCGGACTCCCTGCTCAGCTTGAAGATCAACTTAACTCCCT 6704  
 Db 1065 TGGGACATGAAAGGGGTGCGGACTCCCTGCTCAGCTTGAAGATCAACTTAACTCCCT 1124  
 QY 6705 AGCAGCAGTAGTCTCTTC 6721  
 Db 1125 AGCAGCAGTAGTCTCTTC 1141  
 RESULT 16  
 AAX25659  
 ID AAX25659 standard; cDNA to mRNA; 1948 BP.  
 AC AAX25659;  
 XX  
 DT 21-MAY-1999 (first entry)  
 XX  
 DE Human endogenous retrovirus W clone cl.24.4.  
 XX  
 KW Clone; human endogenous retrovirus; genome; autoimmune disease;  
 KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;  
 KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.  
 XX  
 OS Human endogenous retrovirus.  
 XX  
 PN W09902696-A1.  
 XX  
 PD 21-JAN-1999.  
 XX  
 PF 06-JUL-1998; 98WO-FR001442.  
 XX  
 PR 07-JUL-1997; 97FR-00008815.  
 XX  
 PA (INMR ) BIO MERIEUX.  
 XX  
 PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;  
 XX  
 XX WPI; 1999-120897/10.  
 DR  
 XX  
 PT New nucleic acid sequences from human endogenous retrovirus-W - expressed  
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune  
 PT disease, and abnormal or failed pregnancy.  
 XX  
 XX Claim 1; Page 56-58; 106pp; French.  
 PS  
 CC This sequence represents clone cl.24.4 of the human endogenous retrovirus  
 CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded  
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,  
 CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-  
 CC dependent diabetes and related pathologies) and of abnormal or  
 CC unsuccessful pregnancy and can be used as chromosomal markers for  
 CC susceptibility to these conditions, or proximity markers of genes  
 CC associated with this susceptibility  
 XX  
 SQ Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 0 U; 3 Other;  
 Query Match 9.6%; Score 731; DB 2; Length 1948;

	Best Local Similarity	99.4%; Pred. No. 0;	Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy	5685	AGAGTTTCTATGAGAAATGACAGCGTCCGCGAAATATTGATGSCCCCATCGTATAGAGTCT	5744
Db	823	AGAGTTTCTATGAGAAATGACAGCGTCCGCGAAATATTGATGSCCCCATCGTATAGAGTCT	882
Qy	5745	TTSTAAGGAAACCCCACTTTCATCGGCCACACCCATATGCCCCCGCAACTGCTATCACTC	5804
Db	883	TTSTAAGGAAACCCCACTTTCATCGGCCACACCCATATGCCCCCGCAACTGCTATCACTC	942
Qy	5805	TGCCACTCTTTGCATGATGATGCAATATCTCATTTATGACAGGAAATGATTATCTTAG	5864
Db	943	TGCCACTCTTTGCATGATGCAATATCTCATTTATGACAGGAAATGATTATCTTAG	1002
Qy	5865	TTTGCTCTGGAGGACTTGGAGTCACTGCTGTGTGACCTTACTTACCACCAACTGTTATGTC	5924
Db	1003	TTTGCTCTGGAGGACTTGGAGTCACTGCTGTGTGACCTTACTTACCACCAACTGTTATGTC	1062
Qy	5925	TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	5984
Db	1063	TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	1122
Qy	5985	ACTCACCSGGGTACATGGCACTTACGCCCTACAAAGGACTAGATCTCTCAAACTACA	6044
Db	1123	ACTCACCSGGGTACATGGCACTTACGCCCTACAAAGGACTAGATCTCTCAAACTACA	1182
Qy	6045	TGAACCCCTCCGTACCCATACCTGCCCTGGTAAGGCTATTTTAATACCAACCCTCACTGGGCT	6104
Db	1183	TGAACCCCTCCGTACCCATACCTGCCCTGGTAAGGCTATTTTAATACCAACCCTCACTGGGCT	1242
Qy	6105	CCATGAGGCTCTGGGCCCAAAACCCCTACTAACTGTGTGATATGCTCCCTGAACTTTCA	6164
Db	1243	CCATGAGGCTCTGGGCCCAAAACCCCTACTAACTGTGTGATATGCTCCCTGAACTTTCA	1302
Qy	6165	GCCATATGTTCAATCCCTGTACTGAAACAACTTGAACAACTTCAGACAGAAATAAACAC	6224
Db	1303	GCCATATGTTCAATCCCTGTACTGAAACAACTTGAACAACTTCAGACAGAAATAAACAC	1362
Qy	6225	CACCTCCGTTTTAGTAGGACCTCTGTGTTTCCAAATGCGAAATACCCATACCTCAACCT	6284
Db	1363	CACCTCCGTTTTAGTAGGACCTCTGTGTTTCCAAATGCGAAATACCCATACCTCAACCT	1422
Qy	6285	CACCTGTGTAAAAATTAGCAATACATACACAAACCACTCCCAATGCATCAGGTGGGT	6344
Db	1423	CACCTGTGTAAAAATTAGCAATACATACACAAACCACTCCCAATGCATCAGGTGGGT	1482
Qy	6345	AACCTCTCCACACAAATAGTGTGCTTACCCCTCAGGAATATTTTGTCTGTGGTACCTC	6404
Db	1483	AACCTCTCCACACAAATAGTGTGCTTACCCCTCAGGAATATTTTGTCTGTGGTACCTC	1542
Qy	6405	AGCTATCGTTGTTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTAGTGCC	6464
Db	1543	AGCTATCGTTGTTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTAGTGCC	1602
Qy	6465	CCCYATGRCACCTACACCTGAACAAAGATTATACAGTTATGTCTATCTAAGCCCGCAA	6524
Db	1603	CCCYATGRCACCTACACCTGAACAAAGATTATACAGTTATGTCTATCTAAGCCCGCAA	1662
Qy	6525	CAAAAGAGTACCCATTCTTCTTTTGTATAGGACGAGGAGTGTAGGTGCATAGGTATC	6584
Db	1663	CAAAAGAGTACCCATTCTTCTTTTGTATAGGACGAGGAGTGTAGGTGCATAGGTATC	1722
Qy	6585	TGSCATTTGGCGGTATACAACTCTACTAGTTCTACTACAACTATCTCAAGACTTAA	6644
Db	1723	TGSCATTTGGCGGTATACAACTCTACTAGTTCTACTACAACTATCTCAAGACTTAA	1782
Qy	6645	TGGGGACATGGAAACGGGTCCGCGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCCT	6704
Db	1783	TGGGGACATGGAAACGGGTCCGCGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCCT	1842
Qy	6705	AGCAGCAGTAGTCTCTC 6721	

Db 1843 AGCAGCAGTAGTCTTC 1859

RESULT 17  
AAA59209

ID AAA59209 standard; DNA; 1948 BP.  
XX  
AC AAA59209;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE 5' non coding, 3' pol gene, and 5' env gene of HERV-W from human genome.  
XX  
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;  
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200043521-A2.  
XX  
PD 27-JUL-2000.  
XX  
PF 21-JAN-2000; 2000WO-FR000144.  
XX  
PR 21-JAN-1999; 99FR-00000888.  
XX  
PA (INMR ) BIO MERIEUX.  
XX  
PI Paranhos-Baccala G, Mallet F, Voisset C;  
XX  
DR WPI; 2000-499229/44.  
XX

New nucleic acid from human endogenous retrovirus, useful e.g. for  
diagnosis of autoimmune disease and complications of pregnancy, contains  
at least part of the gag gene.

XX  
PS Disclosure; Page 45; 53pp; French.  
XX

The present sequence represents an endogenous retroviral nucleic acid  
fragment, which is associated with an autoimmune disease, and is  
integrated into the human genome. The fragment is originally derived from  
a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W  
retrovirus is associated with autoimmune disease, failure of pregnancy or  
disorders of pregnancy. The nucleic acid fragment, or proteins derived  
from it, are useful for diagnosis of autoimmune disease (specifically  
multiple sclerosis) and for monitoring pregnancy. The nucleic acid  
fragments may also be used for in situ labelling of isolated chromosomes,  
CC while the transcription product can be used to study or monitor T cell  
proliferation in vitro

XX  
SQ Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 0 U; 3 Other;

Query Match 9.6%; Score 731; DB 3; Length 1948;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAATGCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
DB 823 AGAGTTTCTATGAGAATGCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 882  
QY 5745 TTSTAAGGGAAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804  
DB 883 TTGTAAGGGAAACCCCACTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 942  
QY 5805 TGGCACTCTTTGATGATGCAATAACTCATTTATGGACAGGAAAAATGATTATCTCTAG 5864  
DB 943 TGGCACTCTTTGATGATGCAATAACTCATTTATGGACAGGAAAAATGATTATCTCTAG 1002  
QY 5865 TTGTCCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCCACCCAACTGGTATGTC 5924  
DB 1003 TTGTCCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCCACCCAACTGGTATGTC 1062  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAGAACTAATCTCCCA 5984

Db 1063 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAGAGTAATCTCCCA 1122  
Qy 5985 ACTACCCGGGTACATGGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 6044  
Db 1123 ACTACCCGGGTACATGGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 1182  
Qy 6045 TGAACCCCTCGGTACCCATACGCTCGGTGAAGCCCTATTAATACACCCCTCACTGGGCT 6104  
Db 1183 TGAACCCCTCGGTACCCATACGCTCGGTGAAGCCCTATTAATACACCCCTCACTGGGCT 1242  
Qy 6105 CCATGAGGTCTGGGCCCAAAACCTTAACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAR 6164  
Db 1243 CCATGAGGTCTGGGCCCAAAACCTTAACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAR 1302  
Qy 6165 GCCATATGTTCAATCCCTGTACCTGGAACAACTTGAACAACTTCAGCAGAGAAATAAACAC 6224  
Db 1303 GCCATATGTTCAATCCCTGTACCTGGAACAACTTGAACAACTTCAGCAGAGAAATAAACAC 1362  
Qy 6225 CACTTCCGTTTGTAGTAGGACCTCTTGTTCGAATSTGGAATAACCCATACCTCAAACT 6284  
Db 1363 CACTTCCGTTTGTAGTAGGACCTCTTGTTCGAATSTGGAATAACCCATACCTCAAACT 1422  
Qy 6285 CACTGTGTAAATTTAGCAATCTACATACACAACTCCCAATGCATCAGGTGGGT 6344  
Db 1423 CACTGTGTAAATTTAGCAATCTACATACACAACTCCCAATGCATCAGGTGGGT 1482  
Qy 6345 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTGTGTGACCTC 6404  
Db 1483 AACTCCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTGTGTGACCTC 1542  
Qy 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATATGTCTCTCTCATTTCTAGTGCC 6464  
Db 1543 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATATGTCTCTCTCATTTCTAGTGCC 1602  
Qy 6465 CCATATGRCATCTACCTGAAACAGATTTATACAGTTATGTATATCTTAAGCCCGCAA 6524  
Db 1603 CCATATGRCATCTACCTGAAACAGATTTATACAGTTATGTATATCTTAAGCCCGCAA 1662  
Qy 6525 CAAAGAGTACCCATCTCTTTTGTATAGGAGCAGAGTGTAGGTGCATAGGTAC 6584  
Db 1663 CAAAGAGTACCCATCTCTTTTGTATAGGAGCAGAGTGTAGGTGCATAGGTAC 1722  
Qy 6585 TGGATTTGGCGGTATACAACTCTCTACAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
Db 1723 TGGCATTTGGCGGTATACAACTCTCTACAGTTCTACTACAACTATCTCAAGAACTAAA 1782  
Qy 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCT 6704  
Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCT 1842  
Qy 6705 AGCAGCAGTAGTCCTTC 6721  
Db 1843 AGCAGCAGTAGTCCTTC 1859

RESULT 18  
ABN97948

ID ABN97948 standard; DNA; 2055 BP.

XX AC ABN97948;

XX 01-AUG-2002 (first entry)

XX Human retroviral HERV-7q env coding sequence.

XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;

XX multiple sclerosis; ds.

XX Human endogenous retrovirus.

XX WO9967395-A1.

PD 29-DEC-1999.  
XX 23-JUN-1999; 99MO-FR001513.  
XX 23-JUN-1998; 98FR-00007920.  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX Alliel PM, Perin J, Rieger F;  
XX WPI; 2000-160587/14.  
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used  
XX for diagnosis, treatment and prevention of autoimmune and neurological  
XX diseases.  
XX Claim 3; Page 142-145; 225pp; French.  
XX The present invention relates to new nucleic acid sequences of human  
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.  
XX Regulatory elements associated with HERV-7q may alter expression of other  
XX genes (even remote genes) on the same chromosome, inducing immunological  
XX and/or neurological changes (which may be pathological or protective/  
XX curative). HERV-7q peptides can be used to improve efficiency of the  
XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding  
XX sequences can be used in immunogenic or vaccinating compositions, for  
XX protection against autoimmune diseases, particularly multiple sclerosis.  
XX The peptides may also be used (by sequence comparison) to detect/identify  
XX endogenous retroviruses that are abnormally expressed in cancer,  
XX neuropathologies or other autoimmune diseases. The present sequence was  
XX used to illustrate the invention  
XX Sequence 2055 BP; 576 A; 574 C; 376 G; 529 T; 0 U; 0 Other;  
Qy Query Match 9.6%; Score 731; DB 3; Length 2055;  
Db Best Local Similarity 99.4%; Pred. No. 0;  
Qy Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 5685 AGAGTTTCTATGGAGATGACGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
Db 495 AGAGTTTCTATGGAGATGACGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 554  
Qy 5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCATATGCCGCAACTGCTATCACTC 5804  
Db 555 TTCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCGCAACTGCTATCACTC 614  
Qy 5805 TGCCACTCTTTGTCATGCATGCAAACTACTCATTTTGGACAGAGAAATGATTAATCCTAG 5864  
Db 615 TGCCACTCTTTGTCATGCATGCAAACTACTCATTTTGGACAGAGAAATGATTAATCCTAG 674  
Qy 5865 TTGTCCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCCACCAAACTGGTATGTC 5924  
Db 675 TTGTCCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCCACCAAACTGGTATGTC 734  
Qy 5925 TGATGGGGTGGAGTTCAGAGTCAGGCAAGAGAGAAACATGTAAAGAGTAATCTCCCA 5984  
Db 735 TGATGGGGTGGAGTTCAGAGTTCAGGCAAGAGAGAGAAACATGTAAAGAGTAATCTCCCA 794  
Qy 5985 ACTCACCGGGTACATGGCAGCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044  
Db 795 ACTCACCGGGTACATGGCAGCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 854  
Qy 6045 TGAACCCCTCGGTACCCATACCTGCGCTGGTAAGCCCTATTTAATACACCCCTCACTGGGCT 6104  
Db 855 TGAACCCCTCGGTACCCATACCTGCGCTGGTAAGCCCTATTTAATACACCCCTCACTGGGCT 914  
Qy 6105 CCATGAGGTCTGGGCCCAAAACCTTAACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAR 6164  
Db 915 CCATGAGGTCTGGGCCCAAAACCTTAACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAR 974  
Qy 6165 GCCATATGTTTCAATCCCTGTACCTGGAACAACTTGAACAACTTTCAGCAGAGAAATAAACAC 6224  
Db 975 GCCATATGTTTCAATCCCTGTACCTGGAACAACTTGAACAACTTTCAGCAGAGAAATAAACAC 1034

QY 6225 CACTTCGGTTTATAGGACCTCTTGTTCCTCAATSTGGAATACCATACCTCAAACT 6284  
 Db |||||  
 QY 1035 CACTTCGGTTTATAGGACCTCTTGTTCCTCAATCTGGAATACCATACCTCAAACT 1094  
 Db |||||  
 QY 6285 CACTGTGTAAATTTAGCAATACATACACAACTCCCAATGCAATCAGGTGGT 6344  
 Db |||||  
 QY 1095 CACTGTGTAAATTTAGCAATACATACACAACTCCCAATGCAATCAGGTGGT 1154  
 Db |||||  
 QY 6345 AACTCTCCACACAAATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTC 6404  
 Db |||||  
 QY 1155 AACTCTCCACACAAATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTC 1214  
 Db |||||  
 QY 6405 AGCTATCGTGTGTAATGCTCTCAGATCTATGTGCTCTCTCATCTTAGTGC 6464  
 Db |||||  
 QY 1215 AGCTATCGTGTGTAATGCTCTCAGATCTATGTGCTCTCTCATCTTAGTGC 1274  
 Db |||||  
 QY 6465 CCCTATGCCATCTACACTGAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524  
 Db |||||  
 QY 1275 CCCTATGCCATCTACACTGAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1334  
 Db |||||  
 QY 6525 CAAAGAGTACCACTCTCTCTTTGTATAGGACGAGGTGCTAGGTGCACTAGTAC 6584  
 Db |||||  
 QY 1335 CAAAGAGTACCACTCTCTCTTTGTATAGGACGAGGTGCTAGGTGCACTAGTAC 1394  
 Db |||||  
 QY 6585 TGGCATTGGCGGTATCACAACTCTACTGAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
 Db |||||  
 QY 1395 TGGCATTGGCGGTATCACAACTCTACTGAGTTCTACTACAACTATCTCAAGAACTAAA 1454  
 Db |||||  
 QY 6645 TGGGGACATGGAACGGTCCCGACTCCCTGGTCACTTGCAGATCAAGATCAACTTAACTCCCT 6704  
 Db |||||  
 QY 1455 TGGGGACATGGAACGGTCCCGACTCCCTGGTCACTTGCAGATCAAGATCAACTTAACTCCCT 1514  
 Db |||||  
 QY 6705 AGCAGCAGTAGTCTTTC 6721  
 Db |||||  
 QY 1515 AGCAGCAGTAGTCTTTC 1531  
 Db |||||

RESULT 19  
 AAX25661

ID AAX25661 standard; cDNA to mRNA; 2782 BP.

AC AAX25661;

XX 21-MAY-1999 (first entry)

DT Human endogenous retrovirus W clone cl.PH74.

DE Clone; human endogenous retrovirus; genome; autoimmune disease;

XX multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;

XX disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

XX Human endogenous retrovirus.

OS WO9902696-A1.

XX 21-JAN-1999.

XX 06-JUL-1998; 98WO-FR001442.

XX 07-JUL-1997; 97FR-00008815.

XX (INMR ) BIO MERIEUX.

XX Besene F, Blond J, Bouton O, Mandrand B, Mallet F;

XX WPI; 1999-120897/10.

XX New nucleic acid sequences from human endogenous retrovirus-W - expressed  
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune  
 PT disease, and abnormal or failed pregnancy.

PS Claim 1; Page 60-63; 106pp; French.

XX

This sequence represents clone cl.PH74 of the human endogenous retrovirus  
 (HERV) W genome. The nucleic acids, their fragments or peptides encoded  
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,  
 CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-  
 CC dependent diabetes and related pathologies) and of abnormal or  
 CC unsuccessful pregnancy and can be used as chromosomal markers for  
 CC susceptibility to these conditions, or proximity markers of genes  
 CC associated with this susceptibility

SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 9.6%; Score 731; DB 2; Length 2782;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGAGGTCCTCCGGAATATTTGATGCCCATCGTATAGGAGTCT 5744

Db AGAGTTTCTATGAGAAATGAGGTCCTCCGGAATATTTGATGCCCATCGTATAGGAGTCT 926

QY 5745 TTSTAAGGGAACCCCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804

Db TTCTAAGGGAACCCCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 986

QY 5805 TGGCATTCTTGCATGCATGCAAAATCACTATTATGGACAGGAAAATGATTAATCTCTAG 5864

Db TGGCATTCTTGCATGCATGCAAAATCACTATTATGGACAGGAAAATGATTAATCTCTAG 1046

QY 5865 TTGTCTCTGAGAGTCTGGAGTCACTGCTGTTGGACTTACTTCCACCCCACTGCTATGTC 5924

Db TTGTCTCTGAGAGTCTGGAGTCACTGCTGTTGGACTTACTTCCACCCCACTGCTATGTC 1106

QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA 5984

Db TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA 1166

QY 5985 ACTACCCGGGTACATGGCACTCTAGCCCTCAAAAGACATAGATCTCTCAAAACTACA 6044

Db ACTACCCGGGTACATGGCACTCTAGCCCTCAAAAGACATAGATCTCTCAAAACTACA 1226

QY 6045 TGAACCCCTCCGTACCCATCTGCTGCTGTTAGGCTTATTAATACCACTCTCACTGGCT 6104

Db TGAACCCCTCCGTACCCATCTGCTGCTGTTAGGCTTATTAATACCACTCTCACTGGCT 1286

QY 6105 CCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGTAACCTTCAR 6164

Db CCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGTAACCTTCAG 1346

QY 6165 GCCATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAATCTCAGACAGAAATAACAC 6224

Db GCCATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAATCTCAGACAGAAATAACAC 1406

QY 6225 CACTTCCGTTTATGAGGACCTCTGTTTCCATSTGGAATACCATACCTCAAACT 6284

Db CACTTCCGTTTATGAGGACCTCTGTTTCCATSTGGAATACCATACCTCAAACT 1466

QY 6285 CACTGTGTAAATTTAGCAATACATACACAACCACTCCCAATGATCAGGTGGT 6344

Db CACTGTGTAAATTTAGCAATACATACACAACCACTCCCAATGATCAGGTGGT 1526

QY 6345 AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404

Db AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1586

QY 6405 AGCTATCGTGTGTAATGAGTCTTTCAGAAATCTATGCTCTCTCATCTTAGTGCC 6464

Db AGCTATCGTGTGTAATGAGTCTTTCAGAAATCTATGCTCTCTCATCTTAGTGCC 1646

QY 6465 CCCTATGRCATCTACCTGAAAGATTTATACAGTTATGTGCATATCTAAGCCCGCAA 6524

Db CCCTATGRCATCTACCTGAAAGATTTATACAGTTATGTGCATATCTAAGCCCGCAA 1706

QY 6525 CAAAAGAGTACCCATTCTTCTTTTATAGGACGAGGAGTGTAGGTGCACTAGGTAC 6584

Db	1707	CAAAAGAGTACCCATTCTCTCTTTGTTATAGGACGAGGAGTGCTAGGTGCACCTAGGTAC	1766
Qy	6585	TGGCATTTGGCGGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	6644
Db	1767	TGGCATTTGGCGGGTATCACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	1826
Qy	6645	TGGGGACATGGAAACGGGTGCGCGACTCGCTGGTCACCTTGCAAGATCAACTTAATCCCT	6704
Db	1827	TGGGGACATGGAAACGGGTGCGCGACTCGCTGGTCACCTTGCAAGATCAACTTAATCCCT	1886
Qy	6705	AGCAGCAGTAGTCCTTC	6721
Db	1887	AGCAGCAGTAGTCCTTC	1903

RESULT 20  
AAA59211  
ID AAA59211 standard; DNA; 2782 Bp.

AC AC  
XX AA59211;  
DT 07-NOV-2000 (first entry)

DE 5' non coding, 3' pol gene, and 3' non coding sequences of HERV-W.  
XX  
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;  
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

ca Homo sapiens.  
OS  
XX  
PN WC200043521-A2.

PD	27-JUL-2000.
XX	
PF	21-JAN-2000; 200WO-FR000144.

PR 21-JAN-1999; 99PR-00000888.  
XX  
PA (INMR ) BIO MERIEUX.

PI Paranhos-Baccala G, Mallet F, Voisset C;  
XX  
DR WPI; 2000-499229/44.

PT New nucleic acid from human endogenous retrovirus, useful e.g. for  
PT diagnosis of autoimmune disease and complications of pregnancy, contains  
PT at least part of the gag gene.

PS Disclosure; Page 46-47; 53pp; French.

XX

CC The present sequence represents an endogenous retroviral nucleic acid  
CC factor which is associated with an autoimmune disease and is  
CC present in the plasma of patients with systemic sclerosis.

integrated into the human genome. The fragment is originally derived from a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W retrovirus is associated with autoimmune disease, failure of pregnancy or disorders of pregnancy. The nucleic acid fragment, or proteins derived from it, are useful for diagnosis of autoimmune disease (specifically multiple sclerosis) and for monitoring pregnancy. The nucleic acid fragments may also be used for *in situ* labelling of isolated chromosomes, while the transcription product can be used to study or monitor T cell proliferation *in vitro*.

Query Match	9.6%	Score 731;	DB 3;	Length 2782;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1031;	Conservative	0;	Mismatches	6;
Indels	0;	Gaps	0;	

Qy	5685	AGAGTTTCTATGGAGAAATGACAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT	5744
Db	867	AGAGTTTCTATGGAGAAATGACAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT	926

5745	Qy	TTTAAAGGGAACCCCAACCTTCTCACTCGCCACACCCCATATGCCCCGCAACTGCTATCACTC	5804
927	Db	TTCTAAAGGGAACCCCAACCTTCTCACTCGCCACACCCCATATGCCCCGCAACTGCTATCACTC	986
5805	Qy	TGCACCTCTTTGCATGCATGCACAAATACTCATTTATTCGACAGGAAAAATGATTAACTCCTAG	5864
987	Db	TTGCACCTCTTTGCATGCATGCACAAATACTCATTTATTCGACAGGAAAAATGATTAACTCCTAG	1046
5865	Qy	TTTCTCTGGAGGACTTTGGAGTCACTGCTCTGTGTGGACTTACTTCCACCACAACTGGTATGTC	5924
1047	Db	TTTCTCTGGAGGACTTTGGAGTCACTGCTCTGTGTGGACTTACTTCCACCACAACTGGTATGTC	1106
5925	Qy	TGATGGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	5984
1107	Db	TGATGGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	1166
5985	Qy	ACTCACCSGGTACATGGCACCTCTAGCCCTACAAAGGACTAGACTCTCTCAAAAATTACA	6044
1167	Db	ACTCACCSGGGTACATGGCACCTCTAGCCCTACAAAGGACTAGACTCTCTCAAAAATTACA	1226
6045	Qy	TGAACCCCTCCGTAACCCATACTCGCTGGTGAAGCCTATTTAAATACCAACCTCTCACTGGGCT	6104
1227	Db	TGAACCCCTCCGTAACCCATACTCGCTGGTGAAGCCTATTTAAATACCAACCTCTCACTGGGCT	1286
6105	Qy	CCATGAGGCTCTGGCCCCAAAACCCCTACTAACTGTTGGATATGCTCCCCCTGAACTTCAR	6164
1287	Db	CCATGAGGCTCTGGCCCCAAAACCCCTACTAACTGTTGGATATGCTCCCCCTGAACTTCAG	1346
6165	Qy	GCCATATGTTTCAATCCCTGTACTCGAACAAATGGAAACAACTTCAGCACAGAAATAAACAC	6224
1347	Db	GCCATATGTTTCAATCCCTGTACTCGAACAAATGGAAACAACTTCAGCACAGAAATAAACAC	1406
6225	Qy	CACCTCCGTTTATAGTAGGACCTCTTGTTTCCAAATTTGAAAAATAACCATACCTCAAACCT	6284
1407	Db	CACCTCCGTTTATAGTAGGACCTCTTGTTTCCAAATTTGAAAAATAACCATACCTCAAACCT	1466
6285	Qy	CACCTGTGTAAAAATTTAGCAATACTACATACACAAACCACTCCCAATGCATCAGGTGGGT	6344
1467	Db	CACCTGTGTAAAAATTTAGCAATACTACATACACAAACCACTCCCAATGCATCAGGTGGGT	1526
6345	Qy	AACTCCTCCACACAAATAGTCTGCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	6404
1527	Db	AACTCCTCCACACAAATAGTCTGCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	1586
6405	Qy	AGCCTATCGTTGTTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCC	6464
1587	Db	AGCCTATCGTTGTTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCC	1646
6465	Qy	CCCVATGRCATCTACACTGAACAGAAATTTATACAGTTATGTGCATATCTAAGCCCCGCAA	6524
1647	Db	CCCTATGACCATCTACACTGAAACAGATTTATACAGTTATGTGCATATCTAAGCCCCGCAA	1706
6525	Qy	CAAAAGAGTACCCATCTTCTCTTTGTTATAGAGCAGGAGTGTAGGTGCACCTAGGTAC	6584
1707	Db	CAAAAGAGTACCCATCTTCTCTTTGTTATAGAGCAGGAGTGTAGGTGCACCTAGGTAC	1766
6585	Qy	TGGCATTTGGGGGTATCACAAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	6644
1767	Db	TGGCATTTGGGGGTATCACAAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	1826
6645	Qy	TGGGGACATGGAAACGGGTCCCGCACTCCCTGGTCACTTGGCAAGATCAAACTTAATCCCT	6704
1827	Db	TGGGGACATGGAAACGGGTCCCGCACTCCCTGGTCACTTGGCAAGATCAAACTTAATCCCT	1886
6705	Qy	AGCAGCAGTAGTGCCTTC	6721
1887	Db	AGCAGCAGTAGTGCCTTC	1903

RESULT 21  
AAH20069  
ID AAH20069 standard; DNA; 2782 BP.  
XX



AAH20069;

08-AUG-2001 (first entry)

HERV-W envelope protein G encoding nucleic acid.

Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein; envelope protein; multiple sclerosis-related superantigen; vaccine; surface antigen; transmembrane; multiple sclerosis; neuroprotective; antisense-therapy; autoimmune disorder; ds.

Human endogenous retrovirus.

Key	Location/Qualifiers
5'UTR	1..762
FT	/*tag= a
FT	763..2379
CDS	/*tag= b
FT	/product= "HERV-W envelope protein G"
FT	/transl_except= (pos:730..792,aa:Phe)
FT	/transl_except= (pos:793..795,aa:Thr)
FT	/transl_except= (pos:812..814,aa:Leu)
FT	/transl_except= (pos:818..820,aa:Ser)
FT	/transl_except= (pos:862..864,aa:Tyr)
FT	/transl_except= (pos:865..867,aa:Gln)
FT	/transl_except= (pos:1174..1176,aa:Arg)
FT	/transl_except= (pos:1441..1443,aa:Leu)
FT	/transl_except= (pos:1903..1905,aa:Gln)
FT	/transl_except= (pos:2017..2019,aa:Lys)
FT	/transl_except= (pos:2026..2028,aa:Arg)
FT	/transl_except= (pos:2044..2046,aa:Arg)
FT	/transl_except= (pos:2089..2091,aa:Gln)
FT	/transl_except= (pos:2170..2172,aa:Asn)
3'UTR	2380..2782
FT	/*tag= c

WO200131021-A1.

03-MAY-2001.

30-OCT-2000; 2000WO-EP010659.

28-OCT-1999; 99BP-00402690.

(UYGE-) UNIV GENEVE.

Conrad B, Mach B;

WPI: 2001-316336/33.

P-PSDB; AAB75138.

New human retrovirus HERV-W ENV proteins/peptides having superantigen activity useful for diagnosing and treating multiple sclerosis.

Claim 13; Fig 9; 94pp; English.

On the basis of the PBS t-RNA motif used for the classification of human endogenous retrovirus (HERVs) the full length endogenous provirus which was located on the long arm of human chromosome 7 (7q21-22) has been designated HERV-W. The present invention describes proteins or peptides (I) having superantigen (SAG) activity comprising the ENV protein (ENV) of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I) have neuroprotective activity, and can be used in: vaccines; antisense-therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are useful for diagnosing multiple sclerosis (MS) or HERV-W-associated disorders. (I) are also useful for identifying substances (and optionally recovering) capable of binding to a retroviral superantigen associated with MS, substances capable of blocking SAG activity and substances capable of blocking transcription or translation of HERV-W retroviral superantigen. A protein or peptide derived from (I), modified to be devoid of SAG activity and being capable of generating an immune response against HERV-W retroviral SAG is useful in therapy. Nucleic acid molecules encoding (I) are useful as vaccines against MS. Substances

capable of blocking SAG activity, capable of binding to a retroviral superantigen associated with MS, or capable of blocking transcription or translation of HERV-W retroviral superantigen for use in treating or preventing MS, obtained using (I) are useful for the treatment and CC diagnosis of MS. (I) and nucleic acids encoding them are useful for CC diagnosing autoimmune disease. The present sequence encodes the CC specifically claimed envelope protein of HERV-W designated G

XX

SQ Sequence 2782 BP; 741 A; 768 C; 564 G; 709 T; 0 U; 0 Other;

Query Match	Best Local Similarity	Score	DB	Length
Matches 1031; Conservative	99.4%;	9.6%;	731; DB 5;	Length 2782;
Mismatches	0;	Mismatches	6;	Indels
Gaps	0;	Gaps	0;	Gaps

5685 AGAGTTTCTATGAGAAATGCGGTCCTCCGGAATATTGATGCCCAATATGATAGGAGTCT 5744

867 AGAGTTTCTATGAGAAATGCGGTCCTCCGGAATATTGATGCCCAATATGATAGGAGTCT 926

5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804

927 TTTCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 986

5805 TGCACCTCTTTGCATGCATGCAAAATCACTATTTCGACAGGAAATGATTAATCCCTAG 5864

987 TGCACCTCTTTGCATGCATGCAAAATCACTATTTCGACAGGAAATGATTAATCCCTAG 1046

5865 TTGTCTCGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCCACCAACTGGTATGTC 5924

1047 TTGTCTCGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCCACCAACTGGTATGTC 1106

5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA 5984

1107 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA 1166

5985 ACTCACCGGTACATGGGACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACA 6044

1167 ACTCACCGGTACATGGGACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACA 1226

6045 TGAACCCCTCCGTACCCATCTCGCTGCTGTAAGCCCTATTATATACCACTCTCACTGGCT 6104

1227 TGAACCCCTCCGTACCCATCTCGCTGCTGTAAGCCCTATTATATACCACTCTCACTGGCT 1286

6105 CCATGAGGTCCTGGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGAACTTCAR 6164

1287 CCATGAGGTCCTGGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGAACTTCAG 1346

6165 GCCATATGTTCAATCCCTGTACCTGAAACAATGGACAACTTCAGCAGCAAGTAATACAC 6224

1347 GCCATATGTTCAATCCCTGTACCTGAAACAATGGACAACTTCAGCAGCAAGTAATACAC 1406

6225 CACTTCCGTTTATGAGGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACT 6284

1407 CACTTCCGTTTATGAGGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACT 1466

6285 CACTTGTGTAAATTTAGCAATACATACATACCACTCCCAATGCATCAGGTGGGT 6344

1467 CACTTGTGTAAATTTAGCAATACATACATACCACTCCCAATGCATCAGGTGGGT 1526

6345 AACTCTCCCAACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6404

1527 AACTCTCCCAACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1586

6405 AGCTATCTGTTGTTGAATGGCTCTTTCAGAACTATGCTCTCTCTCACTTCTAGTGCC 6464

1587 AGCTATCTGTTGTTGAATGGCTCTTTCAGAACTATGCTCTCTCTCTCACTTCTAGTGCC 1646

6465 CCCTATGRCATCTACATGAAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 6524

1647 CCCTATGRCATCTACATGAAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAA 1706

6525 CAAAAGAGTACCCTCTCTTTTGTATAGAGCAGGAGTGTAGGTGCATAGGTAC 6584

1707 CAAAAGAGTACCCTCTCTTTTGTATAGAGCAGGAGTGTAGGTGCATAGGTAC 1766

QY 6585 TGGGATTCGGCGTATCACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
 DB 1767 TGGCAATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1826  
 QY 6645 TGGGACATGGAACGGGTGCGCGACTCCCTGTGTCACTTGTGCAAGATCAACTTAACTCCCT 6704  
 DB 1827 TGGGACATGGAACGGGTGCGCGACTCCCTGTGTCACTTGTGCAAGATCAACTTAACTCCCT 1886  
 QY 6705 AGCAGCAGTAGTCCTTC 6721  
 DB 1887 AGCAGCAGTAGTCCTTC 1903

RESULT 22  
 AAD24195  
 ID AAD24195 standard; cDNA; 2930 BP.

XX AAD24195;

XX 07-MAY-2002 (first entry)

XX Human syncytin cDNA.

XX Human; syncytin; preeclampsia; gestational trophoblast disorder;  
 KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;  
 KW envelope gene; human endogenous defective retrovirus; HERV-W; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 930..2546  
 FT CDS /\*tag= a  
 FT /product= "syncytin"  
 FT

PN WO200204678-A2.

XX 17-JAN-2002.

XX 09-JUL-2001; 2001WO-US021719.

XX 07-JUL-2000; 2000US-0216657P.

XX (GEMY ) GENETICS INST INC.

XX Keith JC, McCoy JM, Mi S;

XX WPI; 2002-171727/22.

XX P-PSDB; AAE14540.

PT Identifying a compound for treating a subject with or at risk of  
 PT developing preeclampsia, comprises determining whether the expression or  
 PT activity of syncytin in the cell is modulated in the presence of a test  
 PT compound.

XX Disclosure; Page 39-42; 43pp; English.

XX The invention relates to identifying compounds which are modulators of  
 CC syncytin expression. The syncytin modulators are useful in diagnosis and  
 CC treatment of preeclampsia and gestational trophoblast disorders (e.g.  
 CC choriocarcinoma, hydatiform mole, placental site tumour and missed/  
 CC incomplete abortion). Syncytin is a human gene derived from the envelope  
 CC gene of human endogenous defective retrovirus, HERV-W. The present  
 CC invention is based partly on the discovery that syncytin expression is  
 CC dramatically reduced in preeclampsia, and is also mis-localised to the  
 CC apical syncytiotrophoblast membrane. The present sequence is human  
 CC syncytin cDNA

XX Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;

XX Query Match 9.6%; Score 728; DB 6; Length 2930;

XX Best Local Similarity 98.9%; Pred. No. 0;

XX Matches 1898; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGTAAATAAAGSCACTAGTTTCATGAAA 4861  
 DB 152 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGTAAATAAAGSCACTAGTTTCATGAAA 211  
 QY 4862 TAAATCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAATAAGCCCTGCTTTCCAGG 4921  
 DB 212 TAAATCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAATAAGCCCTGCTTTCCAGG 271  
 QY 4922 CCACAGTAAACCCAGGAGTATCCACAGGCTTACAGAGTATACGATATCACTTACACTCGGCT 4981  
 DB 272 CCACAGTAAACCCAGGAGTATCCACAGGCTTACAGAGTATACGATATCACTTACACTCGGCT 331  
 QY 4982 GAAGGCCACAGTCTCAGGAGAGTCCGAGAAATGAATGAAYACTCAAGAGACATCTAA 5041  
 DB 332 GAAGGCCACAGTCTCAGGAGAGTCCGAGAAATGAATGAAYACTCAAGAGACATCTAA 391  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCATATGGCCCTGCTGTGCTATAGCTTTAAAAA 5101  
 DB 392 AAAAGCAAAACCCAGGAAACCCACCTCATATGGCCCTGCTGTGCTATAGCTTTAAAAA 451  
 QY 5102 GAATCTGCAACTTTCCCAAAAAGCAGGACTTACGCCCATAGAAATGCTGTATGAAGGC 5161  
 DB 452 GAATCTGCAACTTTCCCAAAAAGCAGGACTTACGCCCATAGAAATGCTGTATGAAGGC 511  
 QY 5162 CTTTCATTAACCAATGACTTGTGCTTACCAAGCAGGAGCACTTGTGAGAGACATCAC 5221  
 DB 512 CTTTCATTAACCAATGACTTGTGCTTACCAAGCAGGAGCACTTGTGAGAGACATCAC 571  
 QY 5222 CTCTCTTACCAAAATCAACAAAGTTCTTAAACATTTACAAAGCACTTATCCCTGAGAAGA 5281  
 DB 572 CTCTCTTACCAAAATCAACAAAGTTCTTAAACATTTACAAAGCACTTATCCCTGAGAAGA 631  
 QY 5282 GGGAAAAGAACTATTTCCACCCWGFACATGGTATTTAGTCAAGTCCCTTCTTAATTC 5341  
 DB 632 GGGAAAAGAACTATTTCCACCCCTTGTGACATGGTATTTAGTCAAGTCCCTTCTTAATTC 691  
 QY 5342 CCCATCCCTAGATACATCTGGGAGGACCTTACCCAGTCACTTTTATTTACCCCACTGC 5401  
 DB 692 CCCATCCCTAGATACATCTGGGAGGACCTTACCCAGTCACTTTTATTTACCCCACTGC 751  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCCTGGATCT 5461  
 DB 752 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCCTGGATCT 811  
 QY 5462 GCCAAAGGAACTGAAAATCCAGGAGACAAAGCTAGTATTTCTGTGAACCTCTAGAGA 5521  
 DB 812 GCCAAAGGAACTGAAAATCCAGGAGACAAAGCTAGTATTTCTGTGAACCTCTAGAGA 871  
 QY 5522 TTTGGCCCTGCTCTTCAACCAACACAGGAGGAAAGTAACTAAATCATTAATCCCCA 5581  
 DB 872 TTTGGCCCTGCTCTTCAACCAACACAGGAGGAAAGTAACTAAATCATTAATCCCCA 930  
 QY 5582 TGGSCCTCCCTTATCATATTTTCTKTASTGTSTTTTACCTTCTTCTCCTCTCCTG 5641  
 DB 931 TGGSCCTCCCTTATCATATTTTCTKTASTGTSTTTTACCTTCTTCTCCTCTCCTG 990  
 QY 5642 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCCAGGAGTCTTATGAGAA 5701  
 DB 991 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCCAGGAGTCTTATGAGAA 1050  
 QY 5702 TGCAGCTCCCGAAATATGATGCCCATGATATAGGAGTCTTTTAAAGGAAACCCCCA 5761  
 DB 1051 TGCAGCTCCCGAAATATGATGCCCATGATATAGGAGTCTTTTAAAGGAAACCCCCA 1110  
 QY 5762 CTTTCACCTGCCACACACCCATATGCCCGCAACTGTATCATCTCTGCACTCTTTGATGC 5821  
 DB 1111 CTTTCACCTGCCACACACCCATATGCCCGCAACTGTATCATCTCTGCACTCTTTGATGC 1170  
 QY 5822 ATGCAAAATACATTTATTTGGACAGGAAATGATTAATCTTATGTTCTCTGGAGGACTTG 5881  
 DB 1171 ATGCAAAATACATTTATTTGGACAGGAAATGATTAATCTTATGTTCTCTGGAGGACTTG 1230







CC ameliorating medical conditions in humans and animals, although no  
 CC supporting data is given. Suggested activities include nutritional  
 CC activity, cytokine and cell proliferation/differentiation activity,  
 CC immune stimulating (e.g. as vaccines) or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity,  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, and tumour  
 CC inhibition activity. The polynucleotides are also stated to be useful for  
 CC gene therapy  
 XX  
 SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;  
 Query Match 9.0%; Score 680; DB 2; Length 2946;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 5685 AGAGTTTCTATGAGAAATGACGGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
 DB 1032 AGAGTTTCTATGAGAAATGACGGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 1091  
 QY 5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804  
 DB 1092 TTCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 1151  
 QY 5805 TGCACCTCTTTCATGCATGCAAACTACTATTATTTGGACAGAAAATGATTAATCCTAG 5864  
 DB 1152 TGCCCACTCTTTCATGCATGCAAACTACTATTATTTGGACAGAAAATGATTAATCCTAG 1211  
 QY 5865 TTGTCTCGAGGACTTGGAGTCACTGTCGTGTTGGACTTACTTCAACCAACTGGTATGTC 5924  
 DB 1212 TTGTCTCGAGGACTTGGAGTCACTGTCGTGTTGGACTTACTTCAACCAACTGGTATGTC 1271  
 QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAATCTCCCA 5984  
 DB 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAATCTCCCA 1331  
 QY 5985 ACTACCCGGGTACATGGGACCTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACA 6044  
 DB 1332 ACTACCCGGGTACATGGGACCTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACA 1391  
 QY 6045 TGAACCCCTCGTACCATCTGCTGCTGTAAGCCTATTATACACCCCTCAGTGGCT 6104  
 DB 1392 TGAACCCCTCGTACCATCTGCTGCTGTAAGCCTATTATACACCCCTCAGTGGCT 1451  
 QY 6105 CCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTTGGATATGCTCCCTCGAATTCAR 6164  
 DB 1452 CCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTTGGATATGCTCCCTCGAATTCAG 1511  
 QY 6165 GCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGACAGAAATAAACAC 6224  
 DB 1512 GCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGACAGAAATAAACAC 1571  
 QY 6225 CACTTCGGTTTGTAGGACCTCTGTTTCCAAATGGAATAACCCATACCTCAAACT 6284  
 DB 1572 CACTTCGGTTTGTAGGACCTCTGTTTCCAAATGGAATAACCCATACCTCAAACT 1631  
 QY 6285 CACTCTGTATAAATTTAGCAATACATACACAACCACTCCCAATGCATCAGGTGGGT 6344  
 DB 1632 CACTCTGTATAAATTTAGCAATACATACACAACCACTCCCAATGCATCAGGTGGGT 1691  
 QY 6345 AACTCTCCCAACAATAATGCTGCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC 6404  
 DB 1692 AACTCTCCCAACAATAATGCTGCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC 1751  
 QY 6405 AGCCTATCGTGTGTTGAATGGCTCTTCAAACTATGTCCTCTCAATCTTCTAGTGCC 6464  
 DB 1752 AGCCTATCGTGTGTTGAATGGCTCTTCAAACTATGTCCTCTCAATCTTCTAGTGCC 1811  
 QY 6465 CCCATATGRCCTTACACTGAAACAAGATTATATACAGTTATGTATATCTTAAGCCCGCAA 6524  
 DB 1812 CCCATATGRCCTTACACTGAAACAAGATTATATACAGTTATGTATATCTTAAGCCCGCAA 1871

QY 6525 CAAAGAGTACCATTCTTCTCTTTTATAGGAGCAGAGTGTAGTGCATAGGTAC 6584  
 DB 1872 CAAAGAGTACCATTCTTCTCTTTTATAGGAGCAGAGTGTAGTGCATAGGTAC 1931  
 QY 6585 TGGCATTGGCGGTATACAACTCTTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 6644  
 DB 1932 TGGCATTGGCGGTATACAACTCTTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 1991  
 QY 6645 TGGGACATGAAACGGGTGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAATCCCT 6704  
 DB 1992 TGGGACATGAAACGGGTGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAATCCCT 2051  
 QY 6705 AGCAGCAGTACTCTCTTC 6721  
 DB 2052 AGCAGCAGTACTCTCTTC 2068  
 RESULT 27  
 AAZ59468  
 ID AAZ59468 standard; cDNA; 2946 BP.  
 XX  
 AC AAZ59468;  
 XX  
 DT 11-APR-2000 (first entry)  
 XX  
 DE Human secreted protein AJ172\_2 polynucleotide sequence.  
 XX  
 KW Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;  
 KW placental pathology; metastasis inhibition; nutritional activity;  
 KW immune stimulator; haematopoiesis regulator; tissue growth;  
 KW tumour inhibitor; anti-inflammatory; clone AJ172\_2; ATCC\_98115;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9960020-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 17-MAY-1999; 99WO-US010915.  
 XX  
 PR 18-MAY-1998; 98US-00080478.  
 PR 20-OCT-1998; 98US-00175928.  
 XX  
 PA (GENY ) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Mi S, Treacy M;  
 XX  
 DR WPI; 2000-116311/10.  
 DR P-PSDB; AAY67313.  
 XX  
 PT New polynucleotides encoding secreted cDNA libraries, used to develop  
 products for the diagnosis and treatment of neoplastic disease.  
 XX  
 PS Claim 14; Page 107-108; 149pp; English.  
 XX  
 CC This is the human secreted protein AJ172\_2 nucleotide sequence, obtained  
 from a human adult testes cDNA library. The invention relates to secreted  
 human and murine proteins. The polynucleotides and proteins are predicted  
 to have biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 animals. Detection of the levels of the proteins can be used for the  
 CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents  
 which modulate the expression or function of the proteins may be used for  
 CC treating a neoplastic disease and inhibiting metastasis. Other suggested  
 activities include nutritional activity (e.g. in feeds), cytokine and  
 CC cell proliferation/differentiation activity, immune stimulating (e.g. as  
 vaccines) or suppressing activity, haematopoiesis regulating activity,  
 CC tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 invasion suppressor activity, and tumour inhibition activity. The

CC polynucleotide sequences are also stated to be useful for gene therapy  
 XX  
 SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;  
 Query Match 9.0%; Score 680; DB 3; Length 2946;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGATTTCCTATGGAGATGCGAGTCCCGGGAATATTGATGCCCCATCGTATAGGAGTCT 5744  
 |||||  
 Db 1032 AGATTTCCTATGGAGATGCGAGTCCCGGGAATATTGATGCCCCATCGTATAGGAGTCT 1091  
 |||||

QY 5745 TTSTAAGGAACCCCGACCTTCTAGTCCGACACCATATGCCGCAACTGCTATCACTC 5804  
 |||||  
 Db 1092 TTCTAAGGAACCCCGACCTTCTAGTCCGACACCATATGCCGCAACTGCTATCACTC 1151  
 |||||

QY 5805 TGGCACTCTTTGATGATGATCAATATCTATTTGACAGGAAATGATTAATCTCTAG 5864  
 |||||  
 Db 1152 TGGCACTCTTTGATGATGATCAATATCTATTTGACAGGAAATGATTAATCTCTAG 1211  
 |||||

QY 5865 TTGTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAAACTGGTATGTC 5924  
 |||||  
 Db 1212 TTGTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAAACTGGTATGTC 1271  
 |||||

QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAACTATCTCCCA 5984  
 |||||  
 Db 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAACTATCTCCCA 1331  
 |||||

QY 5985 ACTCACCGGGTACATGCGACCTCTAGCCCTTCAAAAGGACTAGTCTCTCAAACTACA 6044  
 |||||  
 Db 1332 ACTCACCGGGTACATGCGACCTCTAGCCCTTCAAAAGGACTAGTCTCTCAAACTACA 1391  
 |||||

QY 6045 TGAACCTCCGTAACCATCTCGCTGGTAAGCTTATTTAATACCACTCTACTGGGCT 6104  
 |||||  
 Db 1392 TGAACCTCCGTAACCATCTCGCTGGTAAGCTTATTTAATACCACTCTACTGGGCT 1451  
 |||||

QY 6105 CCATGAGGCTCGGCGCAAAACCTTACTACTGTGGATGCTCCCTGCACTTCTCAR 6164  
 |||||  
 Db 1452 CCATGAGGCTCGGCGCAAAACCTTACTACTGTGGATGCTCCCTGCACTTCTCAR 1511  
 |||||

QY 6165 GCCATATGTTTCAATCCCTGCTGCTGAACCAATGGAACAACCTCAGACACAGAAATAACAC 6224  
 |||||  
 Db 1512 GCCATATGTTTCAATCCCTGCTGCTGAACCAATGGAACAACCTCAGACACAGAAATAACAC 1571  
 |||||

QY 6225 CACTCCGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT 6284  
 |||||  
 Db 1572 CACTCCGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT 1631  
 |||||

QY 6285 CACTGTGTAAATTTAGCAATACATACACAACTCCCAATGCAATCAGGTGGGT 6344  
 |||||  
 Db 1632 CACTGTGTAAATTTAGCAATACATACACAACTCCCAATGCAATCAGGTGGGT 1691  
 |||||

QY 6345 AACTCTCCACAAATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACCTC 6404  
 |||||  
 Db 1692 AACTCTCCACAAATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACCTC 1751  
 |||||

QY 6405 AGCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGTGCTCTCTCATTTCTTAGTGCC 6464  
 |||||  
 Db 1752 AGCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGTGCTCTCTCATTTCTTAGTGCC 1811  
 |||||

QY 6465 CCCATGRCATCTACACTGAAACAGATTATACAGTTATGTGATATCTAAGCCCGCAA 6524  
 |||||  
 Db 1812 CCCATGRCATCTACACTGAAACAGATTATACAGTTATGTGATATCTAAGCCCGCAA 1871  
 |||||

QY 6525 CAAAGAGTACCATCTCTCTTTTGTATAGGAGGAGGAGTCTAGGTGCACTAGGTAC 6584  
 |||||  
 Db 1872 CAAAGAGTACCATCTCTCTTTTGTATAGGAGGAGGAGTCTAGGTGCACTAGGTAC 1931  
 |||||

QY 6585 TGGCATTTGGGGTATCACAACTCTACTCAGTTTCTACTACAACTATCTCAAGAACTAAA 6644  
 |||||  
 Db 1932 TGGCATTTGGGGTATCACAACTCTACTCAGTTTCTACTACAACTATCTCAAGAACTAAA 1991  
 |||||

QY 6645 TGGGGACATGGAAACGGGTGCGCGACTCCCTGGTTCACCTTGCAGAGATCAACTTAACCTCCCT 6704  
 |||||

Db 1992 TGGGGACATGGAAACGGGTGCGCGACTCCCTGGTCAAGATCAACTTAACCTCCCT 2051  
 |||||

QY 6705 AGCAGCAGTAGTCTCTTC 6721  
 |||||

Db 2052 AGCAGCAGTAGTCTCTTC 2068  
 |||||

RESULT 28  
 ADC38776  
 ID ADC38776 standard; cDNA; 2946 BP.  
 XX  
 AC ADC38776;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Human cDNA encoding a secreted protein #63.  
 XX  
 KW ss; gene; immune disorder; severe combined immunodeficiency; SCID;  
 KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;  
 KW rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;  
 KW lymphoid cell deficiency; osteoporosis; osteoarthritis;  
 KW peripheral nervous system disease; peripheral neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; coagulation disorder;  
 KW inflammatory disease; systemic inflammatory response syndrome; SIRS;  
 KW ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;  
 KW hypersensitivity; regeneration; neural cell proliferation; fertility;  
 KW tumour; chemokine; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 US2002193567-A1.  
 19-DEC-2002.  
 XX  
 PF 02-APR-2002; 2002US-00114893.  
 XX  
 PR 11-AUG-1995; 95US-00514014.  
 PR 05-APR-1996; 96US-00628364.  
 PR 19-APR-1996; 96US-00635311.  
 PR 07-JUN-1996; 96US-00659224.  
 PR 17-JUN-1996; 96US-00664596.  
 PR 09-JUL-1996; 96US-00677231.  
 PR 26-JUL-1996; 96US-00686878.  
 PR 23-AUG-1996; 96US-00701819.  
 PR 27-SEP-1996; 96US-00721488.  
 PR 27-SEP-1996; 96US-00721798.  
 PR 27-SEP-1996; 96US-00721923.  
 PR 27-SEP-1996; 96US-00721926.  
 PR 25-OCT-1996; 96US-00738367.  
 PR 30-OCT-1996; 96US-00739775.  
 PR 13-JAN-1997; 97US-00783395.  
 PR 10-APR-1997; 97US-00833823.  
 PR 02-JUN-1997; 97US-00867677.  
 PR 05-SEP-1997; 97US-00924838.  
 PR 06-OCT-1999; 99US-00413232.  
 XX  
 (GEMY ) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoey JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;  
 PI Kelleher K;  
 XX  
 WIPI; 2003-657236/62.  
 DR P-PSDB; ADC38777.  
 XX  
 PT Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and  
 PT BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,  
 PT useful for treating e.g. multiple sclerosis and rheumatoid arthritis.  
 XX  
 PS Disclosure; SEQ ID NO 134; 412pp; English.  
 XX  
 CC The invention relates to a protein comprising fully defined AZ302 1



CC protein or BD127 1 6 protein. The polynucleotides are useful for  
 CC expressing recombinant proteins for analysis and are also useful as  
 CC chromosome markers or tags to identify chromosomes or to map related gene  
 CC positions. The proteins are useful as amino acid supplement, carbon  
 CC source, nitrogen source and carbohydrate source. The proteins are useful  
 CC for treating various immune deficiencies and disorders (e.g. severe  
 CC combined immunodeficiency (SCID), autoimmune disorders (e.g. multiple  
 CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic  
 CC reactions (e.g. asthma), myeloid or lymphoid cell deficiencies,  
 CC osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.  
 CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),  
 CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory  
 CC response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease),  
 CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing  
 CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or  
 CC nerve growth or regeneration, for proliferating neural cells and for  
 CC regenerating tumour and brain tissue, for inducing fertility and for  
 CC inhibiting tumour growth. Proteins are also useful as chemokine for  
 CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also  
 CC useful as inhibitors of receptor/ligand interactions. The present  
 CC sequence represents cDNA encoding a human secreted protein.  
 XX

SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

Query Match 9.0%; Score 680; DB 10; Length 2946;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGACGGTCCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744  
 DB 1032 AGAGTTTCTATGAGAAATGACGGTCCCGGAATATTGATGCCCATCGTATAGGAGTCT 1091  
 QY 5745 TTSTAAGGAAACCCCACTTCACTGCCCAACCCCACTATGCGCGCACTGCTATCACTC 5804  
 DB 1092 TTCTAAGGAAACCCCACTTCACTGCCCAACCCCACTATGCGCGCACTGCTATCACTC 1151  
 QY 5805 TGCCACTCTTCGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5864  
 DB 1152 TGCCACTCTTCGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1211  
 QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCAACCAAACTGGTATGTC 5924  
 DB 1212 TTGTCTCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCAACCAAACTGGTATGTC 1271  
 QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGATTAATCTCCA 5984  
 DB 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGATTAATCTCCA 1331  
 QY 5985 ACTCACCGGGTACATGGCCACTCTAGCCCTACAAAGACTAGATCTCTCAAACTACA 6044  
 DB 1332 ACTCACCGGGTACATGGCCACTCTAGCCCTACAAAGACTAGATCTCTCAAACTACA 1391  
 QY 6045 TGAAACCCCTCCGTACCCATCTCGCTGGTAAAGCTTATTAATACCAACCTCACTGGGT 6104  
 DB 1392 TGAAACCCCTCCGTACCCATCTCGCTGGTAAAGCTTATTAATACCAACCTCACTGGGT 1451  
 QY 6105 CCATGAGGTCTCGGCCCAAAACCTACTAATCTGTGGATATGCTCCCTCAACTTCAR 6164  
 DB 1452 CCATGAGGTCTCGGCCCAAAACCTACTAATCTGTGGATATGCTCCCTCAACTTCAG 1511  
 QY 6165 GCATATGTTTCAATCCCTGTACCTGCAATGGACAACTTCAGCAGCAAGAAATAAACAC 6224  
 DB 1512 GCATATGTTTCAATCCCTGTACCTGCAATGGACAACTTCAGCAGCAAGAAATAAACAC 1571  
 QY 6225 CACTTCGGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT 6284  
 DB 1572 CACTTCGGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT 1631  
 QY 6285 CACCTGTGTAATTTAGCAATCTACTACATACACCAACTCTCCAAATGATCAGGTGGT 6344  
 DB 1632 CACCTGTGTAATTTAGCAATCTACTACATACACCAACTCTCCAAATGATCAGGTGGT 1691  
 QY 6345 AACTCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404

DB 1692 AACTCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACCTC 1751  
 QY 6405 AGCCTATCGTTTGTGAATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTGCC 6464  
 DB 1752 AGCCTATCGTTTGTGAATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTGCC 1811  
 QY 6465 CCCATGRCCTATCTACACTGAAACAGATTTATACAGTTTATGTCATATCTAAGCCCCGCA 6524  
 DB 1812 CCCTATGACCACTACACTGAAACAGATTTATACAGTTTATGTCATATCTAAGCCCCGCA 1871  
 QY 6525 CAAAGAGTACCCATTCTTCTCTTTTATAGGAGCAGAGTGTAGGTGCACCTAGGTAC 6584  
 DB 1872 CAAAGAGTACCCATTCTTCTCTTTTATAGGAGCAGAGTGTAGGTGCACCTAGGTAC 1931  
 QY 6585 TGGCATTGGCGGTATCAAACT 6644  
 DB 1932 TGGCATTGGCGGTATCAAACT 1991  
 QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGCTCAGCTTGCAGATCACTTAACCTCCCT 6704  
 DB 1992 TGGGACATGGAACGGGTGCGGACTCCCTGCTCAGCTTGCAGATCACTTAACCTCCCT 2051  
 QY 6705 AGCAGCAGTACTCTCTTC 6721  
 DB 2052 AGCAGCAGTACTCTCTTC 2068

RESULT 29  
 ADU81027  
 ID ADU81027 standard; DNA; 6332 BP.  
 XX AC ADU81027;  
 XX DT 24-FEB-2005 (first entry)  
 XX DE Membrane-fusion protein containing viral vector, pCAGGS/syncytin.  
 XX KW viral vector; enveloped virus; cytoplasmic; membrane-fusion;  
 OS Cytomegalovirus.  
 OS Gallus gallus.  
 OS Oryctolagus cuniculus.  
 OS Simian virus 40.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT enhancer 18..381  
 FT /tag= a  
 FT /note= "Cytomegalovirus early enhancer"  
 FT promoter 382..662  
 FT /tag= b  
 FT /note= "Chicken beta-actin promoter"  
 FT intron 663..1625  
 FT /tag= c  
 FT /note= "Chicken beta-actin intron"  
 FT intron 1626..1670  
 FT /tag= d  
 FT /note= "Rabbit beta-globin intron"  
 FT CDS 1725..3341  
 FT /tag= e  
 FT /note= "Syncytin gene"  
 FT misc\_feature 3132..3338  
 FT /tag= f  
 FT /note= "Cytoplasmic region"  
 FT polyA\_signal 3359..3800  
 FT /tag= g  
 FT /note= "Rabbit beta-globin poly A signal"  
 FT rep\_origin 4002..4344  
 FT /tag= h  
 FT /note= "SV40 replication origin"  
 FT CDS 5341..6201



PR 15-SEP-1999; 99FR-00011793.  
 XX (INRM ) BIO MERIEUX.  
 PA (INRM ) INST NAT SANTE & RECH MEDICALE.  
 XX  
 PI Mallet F, Cosset F, Blond J, Lavillette D, Bouton O, Ruggieri A;  
 XX  
 XX WPI; 2001-226676/23.  
 DR P-PSDB; AAB67652.  
 XX  
 PT Detecting expression of human endogenous retrovirus envelope protein in  
 PT cells of a tissue or culture, from its ability to induce syncytia.  
 XX  
 PS Disclosure; Page 44-45; 57pp; French.  
 XX  
 CC The present sequence encodes a human endogenous retrovirus envelope  
 CC protein. The specification describes a method for detecting expression of  
 CC an envelope protein from a human endogenous retrovirus (HERV), in cells,  
 CC of a tissue or culture. The method comprises detecting syncytia formation  
 CC due to the fusogenic properties of the envelope protein. Envelope  
 CC polypeptides and polynucleotides are used to produce therapeutic or  
 CC prophylactic compositions, particularly for treatment of cancer, to  
 CC correct defects in placental development (or other natural formation of  
 CC other types of syncytia), and to promote adhesion of cells in grafts or  
 CC cellular repair processes. Expression of sequences antisense to the  
 CC polynucleotide are used to prevent formation of syncytia  
 XX  
 SQ Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 U; 0 Other;

Query Match 8.6%; Score 651; DB 5; Length 2781;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1001; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATCAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744  
 Db 866 AGAGTTTCTATGAGAAATCAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 925  
 QY 5745 TTSTAGGGAACCCACCTTCACTGCCCCACACCCCATATGCCCCGGAACCTGCTATCACTC 5804  
 Db 926 TTCTAAGGGAACCCACCTTCACTGCCCCACACCCCATATGCCCCGGAACCTGCTATCACTC 985  
 QY 5805 TGCCACTCTTTCATGCAATGCAAAATCTATTTATGGAGCAAGAAATGATTAATCCTAG 5864  
 Db 986 TGCCACTCTTTCATGCAATGCAAAATCTATTTATGGAGCAAGAAATGATTAATCCTAG 1045  
 QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGTC 5924  
 Db 1046 TTGTCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGTC 1105  
 QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAAATCTCCCA 5984  
 Db 1106 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAAATCTCCCA 1165  
 QY 5985 ACTCACCGGGTACATGGGACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 6044  
 Db 1166 ACTCACCGGGTACATGGGACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 1225  
 QY 6045 TGAAACCCCTCCGTACCCATACCTGCGCTGGTAAAGCTTATTAATACCACTCACTGGGCT 6104  
 Db 1226 TGAAACCCCTCCGTACCCATACCTGCGCTGGTAAAGCTTATTAATACCACTCACTGGGCT 1285  
 QY 6105 CCATGAGGTCCTGGGCCAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTCAR 6164  
 Db 1286 CCATGAGGTCCTGGGCCAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTCAR 1345  
 QY 6165 GCATATGTTTCAATCCCTGTACTGCAACAATGGAACTTTCAGCAGCAAGAAATTAACAC 6224  
 Db 1346 GCATATGTTTCAATCCCTGTACTGCAACAATGGAACTTTCAGCAGCAAGAAATTAACAC 1405  
 QY 6225 CACTTCCGTTTTAGTAGGACCTCTTGTTCCTAATSTGGAATAACCCATACCTCAAACT 6284  
 Db 1406 CACTTCCGTTTTAGTAGGACCTCTTGTTCCTAATCTGGAAATAACCCATACCTCAAACT 1465

QY 6285 CACCTGTGTAATAATTTAGCAATATCTACATACACAAACCAATCCCAATGCATCAGGTGGT 6344  
 Db 1466 CACCTGTGTAATAATTTAGCAATATCTACATACACAAACCAATCCCAATGCATCAGGTGGT 1525  
 QY 6345 AACTCCTCCACACAAATAGTCTGCTTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC 6404  
 Db 1526 AACTCCTCCACACAAATAGTCTGCTTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC 1585  
 QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCC 6464  
 Db 1586 AGCCTATCGTTGTTGAATGGCTCTTCAAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCC 1645  
 QY 6465 CCCYATGCCATCTACACTGAAAGATTTATACAGTTATGTCTATATCTAAGCCCGCAA 6524  
 Db 1646 CCCTATGACCATCTACACTGAAAGATTTATACAGTTATGTCTATATCTAAGCCCGCAA 1705  
 QY 6525 CAAAAGAGTACCCCATTTCTTCTTTTATAGGAGCAGGAGTGTAGGTGCACTAGGTAC 6584  
 Db 1706 CAAAAGAGTACCCCATTTCTTCTTTTATAGGAGCAGGAGTGTAGGTGCACTAGGTAC 1765  
 QY 6585 TGGCATTTGGCGGTATCAAACTCTTCTCAGTTCTTACTACAACTATCTCAAGAACTAAA 6644  
 Db 1766 TGGCATTTGGCGGTATCAAACTCTTCTCAGTTCTTACTACAACTATCTCAAGAACTAAA 1825  
 QY 6645 TGGGACATGGAACGGGTCCGCACTCCCTGCTGACCTTGCAGATCA 6692  
 Db 1826 TGGGACATGGAACGGGTCCGCACTCCCTGCTGACCTTGCAGATCA 1873

Search completed: August 2, 2006, 13:47:07  
 Job time : 4032 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 15:08:55 ; Search time 27430 Seconds  
(without alignments)  
17675.878 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7489.6  
Sequence: 1 caacaatcgggataataacc.....tattaatacttgcartgtg 7582

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank 152:\*  
1: gb\_env:\*  
2: gb\_pat:\*  
3: gb\_ph:\*  
4: gb\_pl:\*  
5: gb\_pr:\*  
6: gb\_ro:\*  
7: gb\_sts:\*  
8: gb\_sy:\*  
9: gb\_un:\*  
10: gb\_vi:\*  
11: gb\_ov:\*  
12: gb\_htg:\*  
13: gb\_in:\*  
14: gb\_om:\*  
15: gb\_ba:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	7489.6	100.0	7582 2 BD196245 Endogenet
2	7489.6	100.0	7582 2 BD267487 Endogenet
3	7489.6	100.0	7582 2 AX000966 Sequence
4	7489.6	100.0	7582 2 AX027480 Sequence
5	6448	86.1	10222 5 AY101582 Homo sapi
6	6448	86.1	10222 5 AY101585 Homo sapi
7	6448	86.1	56093 2 AX329572 Sequence
8	6448	86.1	56093 5 HSAC000064 Human BAC
9	6448	86.1	149194 5 AC007566 Homo sapi
10	6446.4	86.1	10222 5 AY101583 Homo sapi
11	6440	86.0	10222 5 AY101584 Homo sapi
12	6436	85.9	10499 2 BD221808 Nucleic s
13	6436	85.9	10499 2 AR699757 Sequence
14	6436	85.9	10499 2 AX007980 Sequence
15	6346.4	84.7	10229 5 AY101586 Pan trogl
16	6344.8	84.7	10229 5 AY101587 Pan trogl
17	6344.8	84.7	165590 5 AC161288 Pan trogl
18	6344.8	84.7	184675 5 AC145964 Pan trogl

C	19	6344.8	84.7	187270	5	AC161612	AC161612 Pan trogl
	20	6269.6	83.7	10230	5	AY101588	AY101588 Gorilla g
	21	6269.6	83.7	10230	5	AY101589	AY101589 Gorilla g
	22	6124.6	81.8	8523	2	AR612522	AR612522 Sequence
	23	6037.8	80.6	10122	5	AY101590	AY101590 Pongo pyg
	24	6025.8	80.5	10124	5	AY101591	AY101591 Pongo pyg
	25	5956.8	79.5	10246	5	AY101593	AY101593 Hylobates
	26	5950.4	79.4	10248	5	AY101592	AY101592 Hylobates
C	27	5824.2	77.8	158033	5	AC018926	AC018926 Homo sapi
	28	5492	73.3	105989	5	AC008121	AC008121 Homo sapi
	29	5492	73.3	110000	12	AC009727	Continuation (3 of
	30	5380	71.8	169389	12	AC169739	AC169739 Macaca mu
	31	5380	71.8	200397	12	AC144284	AC144284 Macaca mu
C	32	4755.6	63.5	187321	12	AC092510	AC092510 Papio anu
	33	4720.4	63.0	22272	5	AY925148	AY925148 Macaca mu
	34	4302.2	57.4	114621	5	AC005187	AC005187 Homo sapi
	35	4295.8	57.4	46575	12	AC080036	AC080036 Homo sapi
	36	4162	55.6	111140	5	AL592310	AL592310 Human DNA
C	37	3944.6	52.7	183680	5	AC098859	AC098859 Homo sapi
	38	3779.8	50.5	73070	5	AL590143	AL590143 Human DNA
	39	3731.6	49.8	173788	5	AC022555	AC022555 Homo sapi
	40	3731.6	49.8	186540	5	AC090341	AC090341 Homo sapi
C	41	3731.6	49.8	187837	5	AC079065	AC079065 Homo sapi
	42	3716.6	49.6	160883	12	AL158814	AL158814 Homo sapi
C	43	3537.2	47.2	137947	5	HS45114	AL023581 Human DNA
	44	3442.8	46.0	153444	12	AC040948	AC040948 Homo sapi
	45	3442.8	46.0	205035	5	CNS00005	AL049870 Human chr

ALIGNMENTS

RESULT 1	BD196245	7582 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD196245	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.			
DEFINITION	BD196245	unidentified			
ACCESSION	BD196245.1	GI:33006015			
VERSION	JP 2002512530-A/11.	unclassified			
KEYWORDS	unclassified	unclassified sequences.			
SOURCE	1 (bases 1 to 7582)	Beseme, F., Blond, J.L., Bouton, O., Mandrand, B. and Mallet, F.			
ORGANISM	unclassified	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders			
REFERENCE	Patent: JP 2002512530-A 11 23-APR-2002;	BIO MERIEUX			
AUTHORS	OS Unidentified	OS Unidentified			
TITLE	PN JP 2002512530-A/11	PD 23-APR-2002			
JOURNAL	PF 06-JUL-1998 JP 1999508244	PR 07-JUL-1997 FR 97/08815			
COMMENT	PI FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD MANDRAND,	PI FRANCOIS MALLET			
	PC C12N15/48, C07K14/15, C12Q1/68, C07K16/10, G01N33/569 CC	Strandedness: Single;			
	CC Topology: Linear;	CC Endogenetic retroviral sequences, associated with autoimmune diseases			
	CC and/or with pregnancy disorders	CC Key Location/Qualifiers			
	FT source	1. .7582			
FEATURES	Location/Qualifiers	1. .7582			
source	/organism="unidentified"	/mol_type="genomic DNA"			
ORIGIN	/db_xref="taxon:32644"				

Query Match									
Best Local Similarity 100.0%; Score 7489.6; DB 2; Length 7582;									
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CAACAATCGGATATAAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC	60						
Db	1	CAACAATCGGATATAAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC	60						
Qy	61	CTTCCCTTTGTATGGAGCTGTTTCACTGCTATTTCACTCTATTAATCTTGCAACTGCA	120						
Db	61	CTTCCCTTTGTATGGAGCTGTTTCACTGCTATTTCACTCTATTAATCTTGCAACTGCA	120						
Qy	121	CTCTCTCGCTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTCTCACCGTCCACCACTGC	180						
Db	121	CTCTCTCGCTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTCTCACCGTCCACCACTGC	180						
Qy	181	TGTTTGGCACCAACCGCANACCTGCGCTGACTCCCATCCCTCTGGATCCTGCAAGGTGTC	240						
Db	181	TGTTTGGCACCAACCGCANACCTGCGCTGACTCCCATCCCTCTGGATCCTGCAAGGTGTC	240						
Qy	241	CGCTGTGCTCTGATCCAGGARGGCGCCATTTGCGCTCCCAATTTGGGCTAAAGGCTTGC	300						
Db	241	CGCTGTGCTCTGATCCAGGARGGCGCCATTTGCGCTCCCAATTTGGGCTAAAGGCTTGC	300						
Qy	301	CATTGTNCCTGCACGCTTAAGTGCCTGGGTTTGTCTTAATTCAGCTGAACACTANTCACT	360						
Db	301	CATTGTNCCTGCACGCTTAAGTGCCTGGGTTTGTCTTAATTCAGCTGAACACTANTCACT	360						
Qy	361	GGGTTCATGGTTCCTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA	420						
Db	361	GGGTTCATGGTTCCTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA	420						
Qy	421	TGGCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAAGAATAC	480						
Db	421	TGGCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAAGAATAC	480						
Qy	481	GARGCTTGCACATCTTGGAGCGGCTGCTACCBTCTTGAAGTGGTTTCAACCACTC	540						
Db	481	GARGCTTGCACATCTTGGAGCGGCTGCTACCBTCTTGAAGTGGTTTCAACCACTC	540						
Qy	541	TTGGGAGCTCTGTGACGAGGACCCCGGTACATTTTGGCRACCCMSRACGGACATCC	600						
Db	541	TTGGGAGCTCTGTGACGAGGACCCCGGTACATTTTGGCRACCCMSRACGGACATCC	600						
Qy	601	MAAGTGATGGGAAAGCTTCCCGCAAGACAAAACGCCCTCAAGAGTATTCGGARAT	660						
Db	601	MAAGTGATGGGAAAGCTTCCCGCAAGACAAAACGCCCTCAAGAGTATTCGGARAT	660						
Qy	661	TGGGAMCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATTCTTCTGCAGTGCC	720						
Db	661	TGGGAMCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATTCTTCTGCAGTGCC	720						
Qy	721	GCTGTGCACTCTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACTCTTTTGG	780						
Db	721	GCTGTGCACTCTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACTCTTTTGG	780						
Qy	781	TAGAAAAGGCAATGAGTGAAGTGCATAGTACAAAATTTCTTTTCAATTAAGAGACAA	840						
Db	781	TAGAAAAGGCAATGAGTGAAGTGCATAGTACAAAATTTCTTTTCAATTAAGAGACAA	840						
Qy	841	CTCACAAATTTATGTAAGTGTGATTTATGCCCCACAGGAAGCCTTCAGAGTCTACCTCC	900						
Db	841	CTCACAAATTTATGTAAGTGTGATTTATGCCCCACAGGAAGCCTTCAGAGTCTACCTCC	900						
Qy	901	CTATCCAGCATCCCCGACTCTCTTCCCAATTAAGAGACCCCTTCAACCCCAATGG	960						
Db	901	CTATCCAGCATCCCCGACTCTCTTCCCAATTAAGAGACCCCTTCAACCCCAATGG	960						
Qy	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCACAATATTTCCCAAT	1020						
Db	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCACAATATTTCCCAAT	1020						
Qy	1021	TATGACCCCTCCCAAGCAGTGGAGGAAGAAATTCGGCCCGCAGAGTGCATGTGCT	1080						

Db	1021	TATGACCCCTCCCAAGCAGTGGAGGAAGAAATTCGGCCCGCAGAGTGCATGTGCT	1080						
Qy	1081	TTTTTCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATCTCAGATAAATCT	1140						
Db	1081	TTTTTCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATCTCAGATAAATCT	1140						
Qy	1141	GATGGCTATATTTGTGTTTACAGGGTTAGGACAAATCTTTTGATCTCAGATGGAGAT	1200						
Db	1141	GATGGCTATATTTGTGTTTACAGGGTTAGGACAAATCTTTTGATCTCAGATGGAGAT	1200						
Qy	1201	ATATATGTCACTGCTTAAATCAGACACTAACCCCAATGAGAGAGTCCACCATAACTGC	1260						
Db	1201	ATATATGTCACTGCTTAAATCAGACACTAACCCCAATGAGAGAGTCCACCATAACTGC	1260						
Qy	1261	AGCCTGAGRTTGGCGATCTCTGATCTCAGTCAGTCAGTCAATGGATANGATGACAACA	1320						
Db	1261	AGCCTGAGRTTGGCGATCTCTGATCTCAGTCAGTCAGTCAATGGATANGATGACAACA	1320						
Qy	1321	GAAAGGAAGANAATGATTCCCAACAGGCCAGCAGCTCCAGTCTASACCCCTCATTTG	1380						
Db	1321	GAAAGGAAGANAATGATTCCCAACAGGCCAGCAGCTCCAGTCTASACCCCTCATTTG	1380						
Qy	1381	GGGACACAGAAATCAGTAAATCGGAGATTCGTCTGCAGACATTTGCTTAACTTGTGTC	1440						
Db	1381	GGGACACAGAAATCAGTAAATCGGAGATTCGTCTGCAGACATTTGCTTAACTTGTGTC	1440						
Qy	1441	TASAGGACTAAGGAAAACTASGAAAGAAATCTAYGAATTAATCAATGATGCCACATA	1500						
Db	1441	TASAGGACTAAGGAAAACTASGAAAGAAATCTAYGAATTAATCAATGATGCCACATA	1500						
Qy	1501	ACACAGGGAAGGAAGAAATCTTACCTCTTCTGAGAGACTTAAGGGAGGCAATGAG	1560						
Db	1501	ACACAGGGAAGGAAGAAATCTTACCTCTTCTGAGAGACTTAAGGGAGGCAATGAG	1560						
Qy	1561	GAGCGTCCCTCTCTGTCACTGACTCTTCTGAGGCGCAATTAATCTTAAAGCGTAAGTT	1620						
Db	1561	GAGCGTCCCTCTCTGTCACTGACTCTTCTGAGGCGCAATTAATCTTAAAGCGTAAGTT	1620						
Qy	1621	TATCACTCAGTCAGTCGAGACATTTAGAAAAAACTTTCAAAAGTCTGCGCTAGGCCGAG	1680						
Db	1621	TATCACTCAGTCAGTCGAGACATTTAGAAAAAACTTTCAAAAGTCTGCGCTAGGCCGAG	1680						
Qy	1681	CAAACTTTAGAAACCTATTGAACTTGGCAACCTTAAAGGCGCAATTAATAGATCAGGAG	1740						
Db	1681	CAAACTTTAGAAACCTATTGAACTTGGCAACCTTAAAGGCGCAATTAATAGATCAGGAG	1740						
Qy	1741	GAGCAGGCGGAACAGGCAAAACGGGATTTAAAAAAGGCGCACCGCTTTAGTCAATGACCT	1800						
Db	1741	GAGCAGGCGGAACAGGCAAAACGGGATTTAAAAAAGGCGCACCGCTTTAGTCAATGACCT	1800						
Qy	1801	CAGGCAAGTGAATTTGAGGCTCTGAAAAAGGAAAAAGTGGGCAAAATTTGAATGCTTAA	1860						
Db	1801	CAGGCAAGTGAATTTGAGGCTCTGAAAAAGGAAAAAGTGGGCAAAATTTGAATGCTTAA	1860						
Qy	1861	TAGGCTTCTTCCAGTGGGTCTTACAGGACACTTTTAAAGAAATTTCTTCAAGTAGAG	1920						
Db	1861	TAGGCTTCTTCCAGTGGGTCTTACAGGACACTTTTAAAGAAATTTCTTCAAGTAGAG	1920						
Qy	1921	TAAAGCGCCCTTCTGCTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCC	1980						
Db	1921	TAAAGCGCCCTTCTGCTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCC	1980						
Qy	1981	CAGGGAACAAAGGTCTTTTGAAGTCAGAGCCATAACAGATGATCCAGCAGAGACTG	2040						
Db	1981	CAGGGAACAAAGGTCTTTTGAAGTCAGAGCCATAACAGATGATCCAGCAGAGACTG	2040						
Qy	2041	AGGCTGCTGGGCAAGGCCATCCATGCCATCACCTTCAAGAGCCCTGGGTATGCTT	2100						
Db	2041	AGGCTGCTGGGCAAGGCCATCCATGCCATCACCTTCAAGAGCCCTGGGTATGCTT	2100						
Qy	2101	GACCAATTCAGGGCCAGGAAGGTGTCTCTCTGACACTGGTGGCTTCTTCTTACTTACT	2160						

Db 2101 GACCATTTGAGGGCCAGGAAGTTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACT 2160  
Qy 2161 CTTCTGTCCGGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTPAAGAGC 2220  
Db 2161 CTTCTGTCCGGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTPAAGAGC 2220  
Qy 2221 GGCAGTCACTAGATACCTTTTCCAGGCACTAAGTTATGAACCTGGGAGCTTTATTTCTTT 2280  
Db 2221 GGCAGTCACTAGATACCTTTTCCAGGCACTAAGTTATGAACCTGGGAGCTTTATTTCTTT 2280  
Qy 2281 TCACATGCTTTTCTAATATATGCTTTGAAAGCCCCACTACCTTGTGTAGGGAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAATATATGCTTTGAAAGCCCCACTACCTTGTGTAGGGAGAGATTTCTA 2340  
Qy 2341 GCAAAAGCAGGGGCCATATACACCTGAAATAGGAGAGAAACACCCGTTTGTGTGNNC 2400  
Db 2341 GCAAAAGCAGGGGCCATATACACCTGAAATAGGAGAGAAACACCCGTTTGTGTGNNC 2400  
Qy 2401 CCTGCTTGAGGAAGGAATTAATCTGAAGTCTGGGCAACAGAAAGGACAATATGACGAGC 2460  
Db 2401 CCTGCTTGAGGAAGGAATTAATCTGAAGTCTGGGCAACAGAAAGGACAATATGACGAGC 2460  
Qy 2461 CAAAGAAATGCCGTCTCTGTTCAAGTTAAACTAAAGGATTCACCTTCTTCCCTACCAA 2520  
Db 2461 CAAAGAAATGCCGTCTCTGTTCAAGTTAAACTAAAGGATTCACCTTCTTCCCTACCAA 2520  
Qy 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGAGATTCCAAAGATTTGTTAAGGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGAGATTCCAAAGATTTGTTAAGGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTGAGGATTCAGAGATTCAGGATTAATCAATGG 2640  
Db 2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTGAGGATTCAGGATTAATCAATGG 2640  
Qy 2641 GAGGCA CAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAATCAATGG 2700  
Db 2641 GAGGCA CAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAATCAATGG 2700  
Qy 2701 AGGCCGTTGTCTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTCCAAATAC 2760  
Db 2701 AGGCCGTTGTCTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTCCAAATAC 2760  
Qy 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGACCTTMAAGATGCTTCTCTGATCCCTG 2820  
Db 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGACCTTMAAGATGCTTCTCTGATCCCTG 2820  
Qy 2821 TACATCCTGACTCTCAATCTCTGTTGGCTTTTGAAGATCTTCAAACCCCAACATCTCAAC 2880  
Db 2821 TACATCCTGACTCTCAATCTCTGTTGGCTTTTGAAGATCTTCAAACCCCAACATCTCAAC 2880  
Qy 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTTGGCCAGGCAT 2940  
Db 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTTGGCCAGGCAT 2940  
Qy 2941 TAGCCCAAGACTTGGYCATYMTCACTGTGACACTCTGTCTCTGTGTAGTGGATG 3000  
Db 2941 TAGCCCAAGACTTGGYCATYMTCACTGTGACACTCTGTCTCTGTGTAGTGGATG 3000  
Qy 3001 ATTTACTTTTTRGCGCCYRTTTCAGAAACCTTGTGCCATCAAGCCACCCCAAGCRCTCTTWA 3060  
Db 3001 ATTTACTTTTTRGCGCCYRTTTCAGAAACCTTGTGCCATCAAGCCACCCCAAGCRCTCTTWA 3060  
Qy 3061 ATTTCTTCGCYACCTGTGGCTACAWGGTTTCCAAAGSARAGCTCACTCTGCTCACAGC 3120  
Db 3061 ATTTCTTCGCYACCTGTGGCTACAWGGTTTCCAAAGSARAGCTCACTCTGCTCACAGC 3120  
Qy 3121 AGGTTAAATCTTAGGCTTAARATTTATCCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Db 3121 AGGTTAAATCTTAGGCTTAARATTTATCCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Qy 3181 TCCAGCCTTACTGGCTTATCCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTCTTG 3240  
Db 3181 TCCAGCCTTACTGGCTTATCCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTCTTG 3240

Qy 3241 GCRTAAYAGGYTTTCTCCGAATWATGATTTCCAGGTWTGGCRAAATAGCCAGGYCATTA 3300  
Db 3241 GCRTAAYAGGYTTTCTCCGAATWATGATTTCCAGGTWTGGCRAAATAGCCAGGYCATTA 3300  
Qy 3301 WATACASTAATTAAGGAAACTCAGAAAGCCAAATACCATTARTTAAGATGGAAYAMCTGAA 3360  
Db 3301 WATACASTAATTAAGGAAACTCAGAAAGCCAAATACCATTARTTAAGATGGAAYAMCTGAA 3360  
Qy 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
Db 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
Qy 3421 TGCCAA CRGGGCAAGACTTTTSTTYATATRTCAAGAAAAACAGAAAYAGCTCTRGG 3480  
Db 3421 TGCCAA CRGGGCAAGACTTTTSTTYATATRTCAAGAAAAACAGAAAYAGCTCTRGG 3480  
Qy 3481 GTCTCTTACACAGTCCCRAGGGAYGAGCTTGGCAACCTTGGCRVACCTGASTAGGAAAYT 3540  
Db 3481 GTCTCTTACACAGTCCCRAGGGAYGAGCTTGGCAACCTTGGCRVACCTGASTAGGAAAYT 3540  
Qy 3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Db 3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Qy 3601 GTATCTGAAGCAGTTAAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Db 3601 GTATCTGAAGCAGTTAAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Qy 3661 GTGAAYRGCATACTCACTCTAAGGAGACTTGTGGCTGTGACACACYGTTTACTTAA 3720  
Db 3661 GTGAAYRGCATACTCACTCTAAGGAGACTTGTGGCTGTGACACACYGTTTACTTAA 3720  
Qy 3721 TTTCAAGGCTTCTATTACTTGAARGGCCAGTGTGCACTGTGCACTTGTGCACTCTTAAC 3780  
Db 3721 TTTCAAGGCTTCTATTACTTGAARGGCCAGTGTGCACTGTGCACTTGTGCACTCTTAAC 3780  
Qy 3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAAAGATAAAYATAACTGTCAACAARTAAAT 3840  
Db 3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAAAGATAAAYATAACTGTCAACAARTAAAT 3840  
Qy 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTTTAGAGTTCCYTTGACTGATCCYACCTT 3900  
Db 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTTTAGAGTTCCYTTGACTGATCCYACCTT 3900  
Qy 3901 CAACTTGTATCTGATGGAAGTTCCCTTTGTAGAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
Db 3901 CAACTTGTATCTGATGGAAGTTCCCTTTGTAGAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
Qy 3961 AGTGTGCTAGTGAATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
Db 3961 AGTGTGCTAGTGAATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
Qy 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTAGGAGAGRAAAAAAGGGYAAA 4080  
Db 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTAGGAGAGRAAAAAAGGGYAAA 4080  
Qy 4081 TATATATACAGACTCTRTARTATGCTYACCTAGTCTCCATGCCATGMBGCAATAGSAR 4140  
Db 4081 TATATATACAGACTCTRTARTATGCTYACCTAGTCTCCATGCCATGMBGCAATAGSAR 4140  
Qy 4141 AGAAGGGAATTTCCCTAACTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db 4141 AGAAGGGAATTTCCCTAACTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACCTGCTGCGGGTCATCA 4260  
Db 4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACCTGCTGCGGGTCATCA 4260  
Qy 4261 NAAAGGAAGRAAGGGGAAATASAGRGAAAYTGCCCAACAKATATTTAGCMAAAGAGC 4320  
Db 4261 NAAAGGAAGRAAGGGGAAATASAGRGAAAYTGCCCAACAKATATTTAGCMAAAGAGC 4320

Qy	4321	TGCAAGCAGGACCTCCATTAGAAATGCTTTATTTAAATCTTCCCTTAGTATAGGTTAATCC	4380
Db	4321	TGCAAGCAGGACCTCCATTAGAAATGCTTTATTTAAATCTTCCCTTAGTATAGGTTAATCC	4380
Qy	4381	CTTCCGGGAAACAAAGCCCGCTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGG	4440
Db	4381	CTTCCGGGAAACAAAGCCCGCTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGG	4440
Qy	4441	CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAAGAAGGAGAAATCTTTTGCTGCAAC	4500
Db	4441	CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAAGAAGGAGAAATCTTTTGCTGCAAC	4500
Qy	4501	TATCCAATGGAAATTACTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCAC	4560
Db	4501	TATCCAATGGAAATTACTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCAC	4560
Qy	4561	CCATCARATGGCCAAATCATTTATTTACTGGACCHAGGCTTTTCAAACCTATCAAGCARAT	4620
Db	4561	CCATCARATGGCCAAATCATTTATTTACTGGACCHAGGCTTTTCAAACCTATCAAGCARAT	4620
Qy	4621	AKTCAGGGCTGTGAATGTGCCARARAATAATCCCTCGCTVATCGCCAAAGCTCCTTC	4680
Db	4621	AKTCAGGGCTGTGAATGTGCCARARAATAATCCCTCGCTVATCGCCAAAGCTCCTTC	4680
Qy	4681	AGGARAACAARAACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTACCACAAG	4740
Db	4681	AGGARAACAARAACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTACCACAAG	4740
Qy	4741	CCCAAACTCAGGGATTTCACTAGTCTGGGTABARATCTTTCA CGGGTTGGCA	4800
Db	4741	CCCAAACTCAGGGATTTCACTAGTCTGGGTABARATCTTTCA CGGGTTGGCA	4800
Qy	4801	PAGGCTTCCCTGTAGACAGAAAAGCCCAAGAGGTATTAAGGCACCTAGTTTCATGAA	4860
Db	4801	PAGGCTTCCCTGTAGACAGAAAAGCCCAAGAGGTATTAAGGCACCTAGTTTCATGAA	4860
Qy	4861	ATAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAG	4920
Db	4861	ATAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAG	4920
Qy	4921	GCCAAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTACCTGCGCC	4980
Db	4921	GCCAAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTACCTGCGCC	4980
Qy	4981	TGAAGGCCACAGTCTCTCAGGAAAGGTGAGNAANAATGAATGAAYACTCAAGGACATCTA	5040
Db	4981	TGAAGGCCACAGTCTCTCAGGAAAGGTGAGNAANAATGAATGAAYACTCAAGGACATCTA	5040
Qy	5041	AAAAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGTGCTTATAGCCTTAAAA	5100
Db	5041	AAAAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGTGCTTATAGCCTTAAAA	5100
Qy	5101	AGAAATCTGAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG	5160
Db	5101	AGAAATCTGAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG	5160
Qy	5161	CCCTTCATAACCAATGACCTTGTGCTTGA CCAAGACAGCCCACTTAGTTGCAGACATCA	5220
Db	5161	CCCTTCATAACCAATGACCTTGTGCTTGA CCAAGACAGCCCACTTAGTTGCAGACATCA	5220
Qy	5221	CCTCCTTAGCCAAATATCAACAAGTTCTTTAAAAACATTAAGGAAACCTATCCCTGAGAAG	5280
Db	5221	CCTCCTTAGCCAAATATCAACAAGTTCTTTAAAAACATTAAGGAAACCTATCCCTGAGAAG	5280
Qy	5281	AGGGAAAAAAGACTATTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATT	5340
Db	5281	AGGGAAAAAAGACTATTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATT	5340
Qy	5341	CCCATCCCTAGATACATCTCTGGGAGGAGACCCCTACCCAGTCAATTTTATACCCCACTG	5400
Db	5341	CCCATCCCTAGATACATCTCTGGGAGGAGACCCCTACCCAGTCAATTTTATACCCCACTG	5400
Qy	5401	CGGTTAAAGTGGGAGTGGAGTCTTTGGATACATCACTTTAGTCAAAATCTCTGGATAC	5460

[illegible]



6481 ACTGAACAAGATTTATACAGTTATGTTCATATCTAAGCCCCCGCAACAAAAGAGTACCCATT 6540  
6541 CTTCTCTTTTGTATAGAGCAGGAGTGTCTAGTGCATCTAGTACTGCGATTGGCGGTATC 6600  
6541 CTTCTCTTTTGTATAGAGCAGGAGTGTCTAGTGCATCTAGTACTGCGATTGGCGGTATC 6600  
6601 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
6601 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
6661 GTGCCGACATCCCTGGTGCACCTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCT 6720  
6661 GTGCCGACATCCCTGGTGCACCTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCT 6720  
6721 CBAATCGAAGAGCTTTAGACTTGTCTAACCGCTGARAGAGGGGAACTGTTATTTTAA 6780  
6721 CBAATCGAAGAGCTTTAGACTTGTCTAACCGCTGARAGAGGGGAACTGTTATTTTAA 6780  
6781 GGGGAAGAACTGCTTATTTATTTAAATCAATCCGGAATCGTCACTGAGAAGTTAAGAA 6840  
6781 GGGGAAGAACTGCTTATTTATTTAAATCAATCCGGAATCGTCACTGAGAAGTTAAGAA 6840  
6841 ATTCSAGATCGAATACAACTGAKAGCAGARGAGCTTCGAAACACTGGACCCCTGGGGCTC 6900  
6841 ATTCSAGATCGAATACAACTGAKAGCAGARGAGCTTCGAAACACTGGACCCCTGGGGCTC 6900  
6901 CTGAGCCATGAGATGCTGCTGATCTCCCTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
6901 CTGAGCCATGAGATGCTGCTGATCTCCCTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
6961 CTACTCTCTTTGGACCCCTGATCTTTTACCTCTCTTTTAACTTTGCTCTTCCAGAACT 7020  
6961 CTACTCTCTTTGGACCCCTGATCTTTTACCTCTCTTTTAACTTTGCTCTTCCAGAACT 7020  
7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAG 7080  
7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAG 7080  
7081 CCCTGAGCCGGCTGTAGTCCCAAGATCTGATGTTTAAATGACATCAAAAGGACCCCTCT 7140  
7081 CCCTGAGCCGGCTGTAGTCCCAAGATCTGATGTTTAAATGACATCAAAAGGACCCCTCT 7140  
7141 GAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAC 7200  
7141 GAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAC 7200  
7201 GGTSGTCGGCCAACTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACACTGAGA 7260  
7201 GGTSGTCGGCCAACTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACACTGAGA 7260  
7261 GACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCCTAGTGTAGTGTGGGAGGTG 7320  
7261 GACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCCTAGTGTAGTGTGGGAGGTG 7320  
7321 ACCACATCCACCTTTAAACAGGGGCTTGCACTTAGTGTACACCTGACCAATTCAGAGAG 7380  
7321 ACCACATCCACCTTTAAACAGGGGCTTGCACTTAGTGTACACCTGACCAATTCAGAGAG 7380  
7381 CTCACCTAAATGCTAATTAGGCAAGACAGGAGGTAAAGAAATAGCCCAATCATVTTATTC 7440  
7381 CTCACCTAAATGCTAATTAGGCAAGACAGGAGGTAAAGAAATAGCCCAATCATVTTATTC 7440  
7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCCAAGTTCGAGCCCGGCAA 7500  
7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCCAAGTTCGAGCCCGGCAA 7500  
7501 CGGCAACCCCTTTGGGTCCCTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560  
7501 CGGCAACCCCTTTGGGTCCCTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560  
7561 TCTATTAAATCTTGCARCTGCR 7582  
7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 2  
BD267487  
LOCUS

## DEFINITION

7582 bp DNA linear PAT 17-JUL-2003  
Endogenous nucleic fragment associated with an autoimmune disease,  
marking method and reagent.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## ORIGIN

## Query Match

## Best Local Similarity

## Matches 7582;

## Conservative

## Mismatches

## Indels

## Gaps

## Length 7582;

## Pred. No. 0;

## Score 7489.6;

## DB 2;

## Organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

## Location/Qualifiers

## 1..7582

## /organism="Homo sapiens"

## /mol\_type="genomic DNA"

## /db\_xref="taxon:9606"

361	Db	GGGTTCCATTGGTTCTCTTCTGTGTGACCCACGGCTTCTAATTAACAATAAACACTTAAACACTTTACCAACA	420
421	Qy	TGGCCCAAGATTTCCATTTCTTGGAAATCCGTGAGGSCAAACGAACTTCCAGGTTCAGAGAATAC	480
421	Db		
421	Db	TGGCCCAAGATTCATTTCTTGGNAITCCGTGAGGSCAAACGAACTTCCAGGTTCAGAGAATAC	480
481	Qy	GARGCTTGGCCACCATTTTGGAAAGGGGCTGTACCRCTTTGGAAAGTGTTTCAACAATC	540
481	Db		
481	Db	GARGCTTGGCCACCATTTTGGAAAGGGGCTGTACCRCTTTGGAAAGTGTTTCAACCACCATC	540
541	Qy	TTGGGAGCTCTGTGACAAGGACCCCGGTGTACATTTTGGCRACCAACGAGATCC	600
541	Db		
541	Db	TTGGGAGCTCTGTGACAAGGACCCCGGTGTACATTTTGGCRACCAACGAGATCC	600
601	Qy	MAAGTGATGGAAACGTTTCCCGCAAGACAAAACGCCCTTAAGACGTAATTTCTGGARAAT	660
601	Db		
601	Db	MAAGTGATGGAAACGTTTCCCGCAAGACAAAACGCCCTTAAGACGTAATTTCTGGARAAT	660
661	Qy	TGGGAMCAATTTTGACCTCTGACACACTAAGAAAGAAACGACTTATATTCTTCTCGAGTGC	720
661	Db		
661	Db	TGGGAMCAATTTTGACCTCTGACACACTAAGAAAGAAACGACTTATATTCTTCTCGAGTGC	720
721	Qy	GCCTGGCAGCTCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACATCTTTTG	780
721	Db		
721	Db	GCCTGGCAGCTCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACATCTTTTG	780
781	Qy	TAGAAAAGGCCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA	840
781	Db		
781	Db	TAGAAAAGGCCAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA	840
841	Qy	CTCACAATTTATGTAAAAAGTGTGATTTATGCTCTACAGGAAGCTTTCAGAGTCTACCTCC	900
841	Db		
841	Db	CTCACAATTTATGTAAAAAGTGTGATTTATGCTCTACAGGAAGCTTTCAGAGTCTACCTCC	900
901	Qy	CTATCCAGCATCCCGGACTCTTCCCGCAATTAAGAGACCCCTTCAACCCCAATGG	960
901	Db		
901	Db	CTATCCAGCATCCCGGACTCTTCCCGCAATTAAGAGACCCCTTCAACCCCAATGG	960
961	Qy	TCCAAAAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCCCAAT	1020
961	Db		
961	Db	TCCAAAAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCCCAAT	1020
1021	Qy	TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAAATTCGGGCCAGCCAGAGTGCATGTGCT	1080
1021	Db		
1021	Db	TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAAATTCGGGCCAGCCAGAGTGCATGTGCT	1080
1081	Qy	TTTTTCTCCAGACTTAAAGCAATTAATAACAGACTTAGTAAATTTCTCAGATAATYCC	1140
1081	Db		
1081	Db	TTTTTCTCCAGACTTAAAGCAATTAATAACAGACTTAGTAAATTTCTCAGATAATYCC	1140
1141	Qy	GATGGCTATATTGRTGTTTTTCAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGAT	1200
1141	Db		
1141	Db	GATGGCTATATTGRTGTTTTTCAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGAT	1200
1201	Qy	ATATATGTCACTGCTAAATCAGACACTAAACCCCAATCAGAGAAGTGCCACCATTAATCGC	1260
1201	Db		
1201	Db	ATATATGTCACTGCTAAATCAGACACTAAACCCCAATCAGAGAAGTGCCACCATTAATCGC	1260
1261	Qy	AGCTGAGRGTTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAAACA	1320
1261	Db		
1261	Db	AGCTGAGRGTTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAAACA	1320
1321	Qy	GAAGGAAGNAATGATTTCCCAAGGCCAGCGCAGTTCCTCCAGTCTASACCTCAATTG	1380
1321	Db		
1321	Db	GAAGGAAGNAATGATTTCCCAAGGCCAGCGCAGTTCCTCCAGTCTASACCTCAATTG	1380
1381	Qy	GGGACACAGAAATCAGTAACATGGGAGATGGTGCTGCAGACATTTTGTCTAACTTGTGTGC	1440
1381	Db		
1381	Db	GGGACACAGAAATCAGTAACATGGGAGATGGTGCTGCAGACATTTTGTCTAACTTGTGTGC	1440
1441	Qy	TASAAAGGACTAAGGAAACTASGAAGAAATCTATAYGAAATTTACTCAATGATGTCCACCAATA	1500
1441	Db		
1441	Db	TASAAAGGACTAAGGAAACTASGAAGAAATCTATAYGAAATTTACTCAATGATGTCCACCAATA	1500

Qy	1501	ACACGGGGAAGGGAAGAAAACTCTACTGCTTTCTGGAGAGACTAAGGGAGCAATTGAG	1560
Db	1501	ACACGGGGAAGGGAAGAAAACTCTACTGCTTTCTGGAGAGACTAAGGGAGCAATTGAG	1560
Qy	1561	GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAATT	1620
Db	1561	GAAGCGTGCCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAATT	1620
Qy	1621	TATCACTCAGTCAGCTGCGAGACATTTAGAAAAAACTTTCAAAAGTCTCCGCTAGGCCCGGAG	1680
Db	1621	TATCACTCAGTCAGCTGCGAGACATTTAGAAAAAACTTTCAAAAGTCTCCGCTAGGCCCGGAG	1680
Qy	1681	CAAAACTTAGAAAACCTTATGAACCTTGGCAACCTTGAAGGCGCAACCTTAAATCTTAAAGCGTAAATT	1740
Db	1681	CAAAACTTAGAAAACCTTATGAACCTTGGCAACCTTGAAGGCGCAACCTTAAATCTTAAAGCGTAAATT	1740
Qy	1741	GAGCAGGCGGAACAGACAGCAAAAACCGGATTTAAAAAAAAGGCCACCGCTTTTAAATCTTAAAGCGTAAATT	1800
Db	1741	GAGCAGGCGGAACAGACAGCAAAAACCGGATTTAAAAAAAAGGCCACCGCTTTTAAATCTTAAAGCGTAAATT	1800
Qy	1801	CAGGCAAGTGGAATTTTGGAGGCTCTCGAAAAGGGAAGAAAGCTGGGCAAAATGGAATGCTCTAA	1860
Db	1801	CAGGCAAGTGGAATTTTGGAGGCTCTCGAAAAGGGAAGAAAGCTGGGCAAAATGGAATGCTCTAA	1860
Qy	1861	TAGGGCTTGCTTCCAGTGGGGTCTACAAGGACACTTTTAAAAAAGATTTGTCCAAAGTAGAAG	1920
Db	1861	TAGGGCTTGCTTCCAGTGGGGTCTACAAGGACACTTTTAAAAAAGATTTGTCCAAAGTAGAAG	1920
Qy	1921	TAAAGCGCGCCCTTTCGTTCATGCGCCCTTATTTTCAAGGGAATCACTGGAAGGCCCACTGGCCC	1980
Db	1921	TAAAGCGCGCCCTTTCGTTCATGCGCCCTTATTTTCAAGGGAATCACTGGAAGGCCCACTGGCCC	1980
Qy	1981	CAGGGACAAAAGTCTTTTGGAGTTCAGAGCCACTTAACCCAGATCATCTACAGCAGCAGACTG	2040
Db	1981	CAGGGACAAAAGTCTTTTGGAGTTCAGAGCCACTTAACCCAGATCATCTACAGCAGCAGACTG	2040
Qy	2041	AGGGTGCTGGGGCAAGCGCCATCCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTTT	2100
Db	2041	AGGGTGCTGGGGCAAGCGCCATCCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTTT	2100
Qy	2101	GACCATGAGGGCCAGGAAGTTGTCCTCTGGAACCTGGTGCGGTCTTCTTAGCTTTACT	2160
Db	2101	GACCATGAGGGCCAGGAAGTTGTCCTCTGGAACCTGGTGCGGTCTTCTTAGCTTTACT	2160
Qy	2161	CTTCTGTCGGGCAACTGTCTCCAGATCTGTCACTATTCTCAGGGGGTCCNTAAGACG	2220
Db	2161	CTTCTGTCGGGCAACTGTCTCCAGATCTGTCACTATTCTCAGGGGGTCCNTAAGACG	2220
Qy	2221	GGCAGTCACTAGATACCTTTTCCAGCCCACTAAGTTATGAACCTGGGAGCTTTATCTTTT	2280
Db	2221	GGCAGTCACTAGATACCTTTTCCAGCCCACTAAGTTATGAACCTGGGAGCTTTATCTTTT	2280
Qy	2281	TCACATGCTTTTCTAATTATGCTTGAAGGCCCACTACCTTGTGTAGGGAGACATTTCTA	2340
Db	2281	TCACATGCTTTTCTAATTATGCTTGAAGGCCCACTACCTTGTGTAGGGAGACATTTCTA	2340
Qy	2341	GCAAAAGCAGGGGCCATTATACACCTTGAAATAGGAGGAACAACCGCTTTGTTGTTNCC	2400
Db	2341	GCAAAAGCAGGGGCCATTATACACCTTGAAATAGGAGGAACAACCGCTTTGTTGTTNCC	2400
Qy	2401	CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTCGGCCAAAGAGGACAATATGGACGAGC	2460
Db	2401	CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTCGGCCAAAGAGGACAATATGGACGAGC	2460
Qy	2461	CAAAAGATGCCCGTCTCTGTTCAAGTTTAAACTAAGGATTCACCTTCTTTCCCTACCAA	2520
Db	2461	CAAAAGATGCCCGTCTCTGTTCAAGTTTAAACTAAGGATTCACCTTCTTTCCCTACCAA	2520
Qy	2521	GGCAGTACCCCTCAGACCCAGGCCCAACAGGATTCAAAAGATTTGTTAAGCACTTAA	2580
Db	2521	GGCAGTACCCCTCAGACCCAGGCCCAACAGGATTCAAAAGATTTGTTAAGCACTTAA	2580

2581 AAGCCCAAGGCTTAGTAAACCATGCAATACTCCCTGCAGTAATTCGGTAGTGGATTGAG 2640  
Db |||||  
2581 AAGCCCAAGGCTTAGTAAACCATGCAATACTCCCTGCAGTAATTCGGTAGTGGATTGAG 2640  
Qy |||||  
2641 GAGGCAAGAAACCCAGTGGACAGTGGAGGGTTAGTGAAGATCTCAGGATTAATCAATGG 2700  
Db |||||  
2641 GAGGCAAGAAACCCAGTGGACAGTGGAGGGTTAGTGAAGATCTCAGGATTAATCAATGG 2700  
Qy |||||  
2701 AGGCGGTGTCCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCGAAATAC 2760  
Db |||||  
2701 AGGCGGTGTCCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCGAAATAC 2760  
Qy |||||  
2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGGACCTTMAAGGATGCTTCTCTGTCATCCCTG 2820  
Db |||||  
2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGGACCTTMAAGGATGCTTCTCTGTCATCCCTG 2820  
Qy |||||  
2821 TACATCCTGACTCTCAATCTCTGTTTGGCTTTGAAGATCTTCAAAACCCARCATCTCAAC 2880  
Db |||||  
2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCAATCTATTTTGGCCAGGCAT 2940  
Qy |||||  
2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCAATCTATTTTGGCCAGGCAT 2940  
Db |||||  
2941 TAGCCCAAGACTTGAAGYCAATYMTACCTGACACTCTTGTCTTCRGTAGTGGATG 3000  
Qy |||||  
2941 TAGCCCAAGACTTGAAGYCAATYMTACCTGACACTCTTGTCTTCRGTAGTGGATG 3000  
Db |||||  
3001 ATTTACTTTTTRGCGCYRTTCAGAAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMA 3060  
Qy |||||  
3001 ATTTACTTTTTRGCGCYRTTCAGAAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMA 3060  
Db |||||  
3061 ATTTCTCGCYACTGTGGCTACAWGGTTTCCAAAASARAGTCTCARCTCTGCTCAAGC 3120  
Qy |||||  
3061 ATTTCTCGCYACTGTGGCTACAWGGTTTCCAAAASARAGTCTCARCTCTGCTCAAGC 3120  
Db |||||  
3121 AGGTTAAATCTTAGRCCTAARATTTCCAAAGGACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Qy |||||  
3121 AGGTTAAATCTTAGRCCTAARATTTCCAAAGGACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Db |||||  
3181 TCCAGCCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGGRRTTCTTTG 3240  
Qy |||||  
3181 TCCAGCCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGGRRTTCTTTG 3240  
Db |||||  
3241 GRTAAAYAGYTTCTCCGAATWAGATTTCCAGGTTGGCRRAATAGCCAGGYCAITTA 3300  
Qy |||||  
3241 GRTAAAYAGYTTCTCCGAATWAGATTTCCAGGTTGGCRRAATAGCCAGGYCAITTA 3300  
Db |||||  
3301 WATACASTAATTAAGAAACTCAGAAAGCCATACCCATTTTARTTAAGATGGAYAMCTGAA 3360  
Qy |||||  
3301 WATACASTAATTAAGAAACTCAGAAAGCCATACCCATTTTARTTAAGATGGAYAMCTGAA 3360  
Db |||||  
3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTAAACCCAAAGYCCAGGTGTTAAGYT 3420  
Qy |||||  
3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTAAACCCAAAGYCCAGGTGTTAAGYT 3420  
Db |||||  
3421 TGCACACRGGGCAAGACTTTTSTTYATYRTTACAGAAAAAACAAGAAAGCTCTRGA 3480  
Qy |||||  
3421 TGCACACRGGGCAAGACTTTTSTTYATYRTTACAGAAAAAACAAGAAAGCTCTRGA 3480  
Db |||||  
3481 GTCTTTACAGTCCRAGGAYAGCTTGCACACCTVTRGGRYVACTGTAGTAAGGAAYT 3540  
Qy |||||  
3481 GTCTTTACAGTCCRAGGAYAGCTTGCACACCTVTRGGRYVACTGTAGTAAGGAAYT 3540  
Db |||||  
3541 GATGTAGTGGCAAGGTTGRCYTCATTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
Qy |||||  
3541 GATGTAGTGGCAAGGTTGRCYTCATTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
Db |||||  
3601 GTATCTGAAGCAGTTAAATATATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Qy |||||  
3601 GTATCTGAAGCAGTTAAATATATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Db |||||  
3661 GTGAAYRGCACTACTGCTAAAGGAGACTTGTGGCTGTGCAGACAACTGTTTACTTAAA 3720  
Qy |||||

3661 GTGAAYRGCACTACTGCTAAAGGAGACTTGTGGCTGTGCAGACAACTGTTTACTTAAA 3720  
Db |||||  
3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAAC 3780  
Qy |||||  
3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAAC 3780  
Db |||||  
3781 CCAGYCNCAATTTCTCCAGCAATGAAGAAAGATARAAYATAAATGTCAACAARTAAAT 3840  
Qy |||||  
3781 CCAGYCNCAATTTCTCCAGCAATGAAGAAAGATARAAYATAAATGTCAACAARTAAAT 3840  
Db |||||  
3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGARGTCCYTTGACTGATCCYACCTT 3900  
Qy |||||  
3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGARGTCCYTTGACTGATCCYACCTT 3900  
Db |||||  
3901 CAACTTGTATCTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
Qy |||||  
3901 CAACTTGTATCTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
Db |||||  
3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTCTYA 4020  
Qy |||||  
3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTCTYA 4020  
Db |||||  
4021 GCTRCAGAACTAATAGCCCTCACTGKGGCACTAGNAATTTAGGAGAAAGAAAGGYYAAA 4080  
Qy |||||  
4021 GCTRCAGAACTAATAGCCCTCACTGKGGCACTAGNAATTTAGGAGAAAGAAAGGYYAAA 4080  
Db |||||  
4081 TATATATACAGACTCTTRATATGCTVACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
Qy |||||  
4081 TATATATACAGACTCTTRATATGCTVACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
Db |||||  
4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATATCMACATCAGAAAGCCATTAGGAR 4200  
Qy |||||  
4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATATCMACATCAGAAAGCCATTAGGAR 4200  
Db |||||  
4201 ATTTATTTAYTGGCWGTACAGAAACCTTARAGAGTGGWAGTCTTACACTGCTGCGGTCTCA 4260  
Qy |||||  
4201 ATTTATTTAYTGGCWGTACAGAAACCTTARAGAGTGGWAGTCTTACACTGCTGCGGTCTCA 4260  
Db |||||  
4261 NAAAGAAAGRAAGAAATASAGRAAYTGCACAGCAKATATTTGAAGCAAAAGAGC 4320  
Qy |||||  
4261 NAAAGAAAGRAAGAAATASAGRAAYTGCACAGCAKATATTTGAAGCAAAAGAGC 4320  
Db |||||  
4321 TGCAGGCAAGGACCTTCACTTATAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAACTC 4380  
Qy |||||  
4321 TGCAGGCAAGGACCTTCACTTATAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAACTC 4380  
Db |||||  
4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGGAG 4440  
Qy |||||  
4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGGAG 4440  
Db |||||  
4441 CAGTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGAAATACTTTTCCCTGCAAC 4500  
Qy |||||  
4441 CAGTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGAAATACTTTTCCCTGCAAC 4500  
Db |||||  
4501 TATCAATGGAATTTACTTTAAACCTTCACTCAAACTTCACTTAGGCATCGATAGCAC 4560  
Qy |||||  
4501 TATCAATGGAATTTACTTTAAACCTTCACTCAAACTTCACTTAGGCATCGATAGCAC 4560  
Db |||||  
4561 CCATCARATGGCCAAATCATTTATTTATCTGGAACAGGCTTTTCAAACTATCAAGCARAT 4620  
Qy |||||  
4561 CCATCARATGGCCAAATCATTTATTTATCTGGAACAGGCTTTTCAAACTATCAAGCARAT 4620  
Db |||||  
4621 AKTCAAGGCTCTGGAATGTGCCARAAATTAATCCCTGCTCTATCGCAGAGCTCTTTC 4680  
Qy |||||  
4621 AKTCAAGGCTCTGGAATGTGCCARAAATTAATCCCTGCTCTATCGCAGAGCTCTTTC 4680  
Db |||||  
4681 AGGAAACAAARAAACAGGCCATTCCTGRRARAACTGGCAACTGATTTTACCACAG 4740  
Qy |||||  
4681 AGGAAACAAARAAACAGGCCATTCCTGRRARAACTGGCAACTGATTTTACCACAG 4740  
Db |||||  
4741 CCCAAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATATTTTCACGGTTGGCA 4800  
Qy |||||

Db 4741 CCCAAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATACATTTTCACGGGTTGGCA 4800  
Qy 4801 RAGGCTTTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCACATAGTTTCATGAA 4860  
Db 4801 RAGGCTTTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCACATAGTTTCATGAA 4860  
Qy 4861 ATAATTTCCAGANTTCGGACTTCCCGAGGCTTACAGAGTGCAATAGAGCCCTGCTTTCCAG 4920  
Db 4861 ATAATTTCCAGANTTCGGACTTCCCGAGGCTTACAGAGTGCAATAGAGCCCTGCTTTCCAG 4920  
Qy 4921 GCCACAGTAACCCAGGGAGTATCCAGAGGCTTGGGTATACGATATCACTTACACTGCGCC 4980  
Db 4921 GCCACAGTAACCCAGGGAGTATCCAGAGGCTTGGGTATACGATATCACTTACACTGCGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCTAGGGAGGTTCAGAGAAATGAATGAATACTCAAAGGACATCTA 5040  
Db 4981 TGAAGGCCACAGTCTCTAGGGAGGTTCAGAGAAATGAATGAATACTCAAAGGACATCTA 5040  
Qy 5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGVTCTGTTGCCCTATAGCCCTTAAAA 5100  
Db 5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGVTCTGTTGCCCTATAGCCCTTAAAA 5100  
Qy 5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
Db 5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
Qy 5161 CCCTTCATAAACCATGACTTGTGCTTGACCCOAGACAGCCAACTTAGTTGGACATCA 5220  
Db 5161 CCCTTCATAAACCATGACTTGTGCTTGACCCOAGACAGCCAACTTAGTTGGACATCA 5220  
Qy 5221 CCTCCTTAGCCAAATATCAACAGTTCTTAAACATTAACAGAAACCTATCCCTCAGAGG 5280  
Db 5221 CCTCCTTAGCCAAATATCAACAGTTCTTAAACATTAACAGAAACCTATCCCTCAGAGG 5280  
Qy 5281 AGGGAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCAATT 5340  
Db 5281 AGGGAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCAATT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCGGAGGACCCCTACCCAGTCAATTTATVTAACCCAACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCGGAGGACCCCTACCCAGTCAATTTATVTAACCCAACTG 5400  
Qy 5401 CGGTTAAAGTGCTGAGTGGAGTCTTGGATACATCACACTTGAAGTCAAACTCCTGGATAC 5460  
Db 5401 CGGTTAAAGTGCTGAGTGGAGTCTTGGATACATCACACTTGAAGTCAAACTCCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACTCGAAATCCAGGAGCAACGCTAGCTATTTCCTGTGAACCTCTAGAGG 5520  
Db 5461 TGCCAAAGGAACTCGAAATCCAGGAGCAACGCTAGCTATTTCCTGTGAACCTCTAGAGG 5520  
Qy 5521 ATTTGGGCTGCTCTTCAAACAAACCAAGAGGAAAGTAACTAAATCATAAATCCCCCC 5580  
Db 5521 ATTTGGGCTGCTCTTCAAACAAACCAAGAGGAAAGTAACTAAATCATAAATCCCCCC 5580  
Qy 5581 ATGSSCCTCCCTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCACT 5640  
Db 5581 ATGSSCCTCCCTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCGCTGTATGACCAAGTACTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCCATGCGCTGTATGACCAAGTACTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGCAAGCTCCCGGAAATATTGATGCCCCCATCGTATAGAGTCTTTTSTAAGGAAACCCCC 5760  
Db 5701 ATGCAAGCTCCCGGAAATATTGATGCCCCCATCGTATAGAGTCTTTTSTAAGGAAACCCCC 5760  
Qy 5761 ACCTTCACCTGCCACACCATATGCCCGCACTGCTATCACTCTGCGCACTCTTTGCAATG 5820  
Db 5761 ACCTTCACCTGCCACACCATATGCCCGCACTGCTATCACTCTGCGCACTCTTTGCAATG 5820  
Qy 5821 CATGAAATACCTATTATTGGACAGGAAATGATTAATCTTAGTTGCTCGGAGACTT 5880  
Db 5821 CATGAAATACCTATTATTGGACAGGAAATGATTAATCTTAGTTGCTCGGAGACTT 5880

Qy 5881 GGAGTCACTGCTCTGTGGACTTACTTCAACCCAAACTGGTATGTCTGATGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGCTCTGTGGACTTACTTCAACCCAAACTGGTATGTCTGATGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACAT 6000  
Qy 6001 GGCACTCTAGCCCTCAAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 6060  
Db 6001 GGCACTCTAGCCCTCAAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 6060  
Qy 6061 CATACTCGCTGGTAAAGCTATTAAATACCAACCTCTACTGGGCTCCATGAGGTCTCGGCC 6120  
Db 6061 CATACTCGCTGGTAAAGCTATTAAATACCAACCTCTACTGGGCTCCATGAGGTCTCGGCC 6120  
Qy 6121 CAAAACCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
Db 6121 CAAAACCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
Qy 6181 CCTGTACTGAACTATGGAACAACTTCAGCACAGAAATAAACACACCTTCGGTTTAGTA 6240  
Db 6181 CCTGTACTGAACTATGGAACAACTTCAGCACAGAAATAAACACACCTTCGGTTTAGTA 6240  
Qy 6241 GGACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAAAACCTCACCTGTGTAATAAT 6300  
Db 6241 GGACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAAAACCTCACCTGTGTAATAAT 6300  
Qy 6301 AGCAATACTACATACACAAACCACTCCCAATGCACTCAGGTGGGTAACTCTCCCAACAAA 6360  
Db 6301 AGCAATACTACATACACAAACCACTCCCAATGCACTCAGGTGGGTAACTCTCCCAACAAA 6360  
Qy 6361 ATAGTCTGCCCTACCTCAGGAATAATTTTGTCTGTGGTACTCAGCCTATGCTGTTGTTG 6420  
Db 6361 ATAGTCTGCCCTACCTCAGGAATAATTTTGTCTGTGGTACTCAGCCTATGCTGTTGTTG 6420  
Qy 6421 AATGGCTCTCAGAACTATGCTGCTCTCTCATCTTAGTGGCCCTCATGRCATCTAC 5480  
Db 6421 AATGGCTCTCAGAACTATGCTGCTCTCTCATCTTAGTGGCCCTCATGRCATCTAC 5480  
Qy 6481 ACTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
Db 6481 ACTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
Qy 6541 CTTCTTTTGTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCAATTGGCGGTATC 6600  
Db 6541 CTTCTTTTGTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCAATTGGCGGTATC 6600  
Qy 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Db 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Qy 6661 GTGCGGACTCCCTGGTCACTTGAAGTCACTTAACTCCCTAGCAGCAGTAGTCTCTT 6720  
Db 6661 GTGCGGACTCCCTGGTCACTTGAAGTCACTTAACTCCCTAGCAGCAGTAGTCTCTT 6720  
Qy 6721 CRAAATCGAAGAGCTTTAGACTYGTAAACCGCTGABAGAGGGGAACTGTTTATTTTAA 6780  
Db 6721 CRAAATCGAAGAGCTTTAGACTYGTAAACCGCTGABAGAGGGGAACTGTTTATTTTAA 6780  
Qy 6781 GGGGAGAGTCTGTTTATTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTBAAGAA 6840  
Db 6781 GGGGAGAGTCTGTTTATTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTBAAGAA 6840  
Qy 6841 ATTSAGATCGAATACAACTGAKAGCAGAGAGCTTCCGAAACACTGGACCCCTGGGCGCTC 6900  
Db 6841 ATTSAGATCGAATACAACTGAKAGCAGAGAGCTTCCGAAACACTGGACCCCTGGGCGCTC 6900  
Qy 6901 CTCAGCCATGATGCTCCCTGGATTCCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Db 6901 CTCAGCCATGATGCTCCCTGGATTCCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG 6960

```
QY 6961 CTACTCCTCTTTGGACCCCTGTATCTTTTRACCTCCTTGTAACTTTGTCTCTTCCAGAAATC 7020
Db 6961 CTACTCCTCTTTGGACCCCTGTATCTTTTRACCTCCTTGTAACTTTGTCTCTTCCAGAAATC 7020
QY 7021 GAAGCTGTAAACTCAAAATGAGCCCAAGATCGAGTCCAAAGACTAAGATCTACCCGAGA 7080
Db 7021 GAAGCTGTAAACTCAAAATGAGCCCAAGATCGAGTCCAAAGACTAAGATCTACCCGAGA 7080
QY 7081 CCCCTGGACGGGCTGTAGCCACGATCTGATGTTAATGACATCAAGGACCCCTCCT 7140
Db 7081 CCCCTGGACGGGCTGTAGCCACGATCTGATGTTAATGACATCAAGGACCCCTCCT 7140
QY 7141 GAGGAATCTCAGTGCACAACTCTACTAGCCGCCCAATTCAGCAGGAAGCAGTTAGAGC 7200
Db 7141 GAGGAATCTCAGTGCACAACTCTACTAGCCGCCCAATTCAGCAGGAAGCAGTTAGAGC 7200
QY 7201 GGTSGTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGA 7260
Db 7201 GGTSGTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGA 7260
QY 7261 GACGAGACTAGCTGGATTTCTAGGCTGATTAAGATCCYTAAGCCTAGSTGGGAAGGTG 7320
Db 7261 GACGAGACTAGCTGGATTTCTAGGCTGATTAAGATCCYTAAGCCTAGSTGGGAAGGTG 7320
QY 7321 ACCACATCCACCTTTAAACACAGGGGCTTGCACCTTAGYTCACACCTGACCAATCAGAGAG 7380
Db 7321 ACCACATCCACCTTTAAACACAGGGGCTTGCACCTTAGYTCACACCTGACCAATCAGAGAG 7380
QY 7381 CTCACATAAATGCTAATTAGGCAAGACAGGAGTTAAGAAATAGCCAATCATYATATGC 7440
Db 7381 CTCACATAAATGCTAATTAGGCAAGACAGGAGTTAAGAAATAGCCAATCATYATATGC 7440
QY 7441 MTGAGAGCAGCAGGAGGACAATGATCGGATATAAACCCAAAGTTCGAGCCGGCAA 7500
Db 7441 MTGAGAGCAGCAGGAGGACAATGATCGGATATAAACCCAAAGTTCGAGCCGGCAA 7500
QY 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTCAATGCTATTTCAC 7560
Db 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTCAATGCTATTTCAC 7560
QY 7561 TCTATTAAATCTTGCACTGCR 7582
Db 7561 TCTATTAAATCTTGCACTGCR 7582

RESULT 3
LOCUS AX000966 7582 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 11 from Patent WO9902696.
ACCESSION AX000966
VERSION AX000966.1 GI:7241208
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 7582)
AUTHORS Beseme, F. and Blond, J.
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS
JOURNAL Patent: WO 902696-A 11 21-JAN-1999;
BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES
source
1..7582
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 7489.6; DB 2; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAATCGGGATATAAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60
```

Db 1081 TTTYVTCTCCAGACTTAAGCAAAATAAAAAACAGACTTAGGTAAATTTCTCAGATAAAYCCT 1140  
Qy 1141 GATGCTATATGRTGTTTTTCAAGGGTTTAGACAATTTCTTTGATCTGACATGGAGAGAT 1200  
Db 1141 GATGCTATATGRTGTTTTTCAAGGGTTTAGACAATTTCTTTGATCTGACATGGAGAGAT 1200  
Qy 1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATAGAGAAGTGCACCAATACCTGC 1260  
Db 1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATAGAGAAGTGCACCAATACCTGC 1260  
Qy 1261 AGCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGATGANGATGACAACA 1320  
Db 1261 AGCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGATGANGATGACAACA 1320  
Qy 1321 GAAGAAAGANAATGATTTCCCAAGGCCAGCARGCAGTTCCTCAGTCTASACCTCATTTG 1380  
Db 1321 GAAGAAAGANAATGATTTCCCAAGGCCAGCARGCAGTTCCTCAGTCTASACCTCATTTG 1380  
Qy 1381 GGGCACACGAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTTAAGTCTGTGTC 1440  
Db 1381 GGGCACACGAATCAGTAACATGGGAGATTGGTGCTGCAGACATTTGCTTAAGTCTGTGTC 1440  
Qy 1441 TASAGGACTAAGGAAAACCTASGAAGAAATCTAYGAATTTACTCAATGATGTCCACCATA 1500  
Db 1441 TASAGGACTAAGGAAAACCTASGAAGAAATCTAYGAATTTACTCAATGATGTCCACCATA 1500  
Qy 1501 ACAGGGGAAGGGAAGAAAATCTCTGCTCTTCTGGAGAGACTAAGGGAGGCAATTTGAG 1560  
Db 1501 ACAGGGGAAGGGAAGAAAATCTCTGCTCTTCTGGAGAGACTAAGGGAGGCAATTTGAG 1560  
Qy 1561 GAAGGCTGCTCTCTGTCACTGACTCTCTGCAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
Db 1561 GAAGGCTGCTCTCTGTCACTGACTCTCTGCAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
Qy 1621 TATCACTCAGTCAGTCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
Db 1621 TATCACTCAGTCAGTCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
Qy 1681 CAAAACCTTAGAAACCTATTGAACCTTGGCAACTGCGTCTTAAATAGAGATCAGGAG 1740  
Db 1681 CAAAACCTTAGAAACCTATTGAACCTTGGCAACTGCGTCTTAAATAGAGATCAGGAG 1740  
Qy 1741 GAGCAGGCGGACAGCAAAACGGGATTAABAAAAGGCCACCGCTTTAGTCATCACCT 1800  
Db 1741 GAGCAGGCGGACAGCAAAACGGGATTAABAAAAGGCCACCGCTTTAGTCATCACCT 1800  
Qy 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAGGAAAAGCTGGGCAAAATTTGAATGCCTAA 1860  
Db 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAGGAAAAGCTGGGCAAAATTTGAATGCCTAA 1860  
Qy 1861 TAGGGCTTGCTTCCAGTGGGCTTACAGGAACATTTTAAAGAGATTTGTCCAAGTAGAAG 1920  
Db 1861 TAGGGCTTGCTTCCAGTGGGCTTACAGGAACATTTTAAAGAGATTTGTCCAAGTAGAAG 1920  
Qy 1921 TAAGCGCCCTTGTCTCATGCCCTTATTTCAAGGGATCACTGGAGGCCCATCTGCC 1980  
Db 1921 TAAGCGCCCTTGTCTCATGCCCTTATTTCAAGGGATCACTGGAGGCCCATCTGCC 1980  
Qy 1981 CAGGGACAAGGCTCTTTTGAAGTCAGAGCCACTAACAGATGATCCAGCAGCAGACTG 2040  
Db 1981 CAGGGACAAGGCTCTTTTGAAGTCAGAGCCACTAACAGATGATCCAGCAGCAGACTG 2040  
Qy 2041 AGGGTCCCTGGGCAAGGCCATCCCATGCCATCACTTCAAGAGCCCTGGGATGCTT 2100  
Db 2041 AGGGTCCCTGGGCAAGGCCATCCCATGCCATCACTTCAAGAGCCCTGGGATGCTT 2100  
Qy 2101 GACCATGAGGCGCAGGAGGTTGCTCTCTGACACTGGTGGCTTCTTCTTACT 2160  
Db 2101 GACCATGAGGCGCAGGAGGTTGCTCTCTGACACTGGTGGCTTCTTCTTACT 2160  
Qy 2161 CTTCTGTCGGGACAACTCTCTCCAGATCTGTCACTATTTCTGAGGGGTCCTNTAAGAG 2220  
Db 2161 CTTCTGTCGGGACAACTCTCTCCAGATCTGTCACTATTTCTGAGGGGTCCTNTAAGAG 2220

Qy 2221 GGCAGTCACTAGATACCTTTTCCAGCCCACTAAGTTATGAAGTGGGAGCTTTATCTTT 2280  
Db 2221 GGCAGTCACTAGATACCTTTTCCAGCCCACTAAGTTATGAAGTGGGAGCTTTATCTTT 2280  
Qy 2281 TCACATGCTTTTCTAAATTTATGTTGAAAGCCCACTACTCTTGTAGGAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAAATTTATGTTGAAAGCCCACTACTCTTGTAGGAGAGACATTTCTA 2340  
Qy 2341 GCAAAAGCAGGGCCATTTATACCTGAAATAGGAGAAGAAACCCGTTTGTGTNCC 2400  
Db 2341 GCAAAAGCAGGGCCATTTATACCTGAAATAGGAGAAGAAACCCGTTTGTGTNCC 2400  
Qy 2401 CCTGTTGAGAAAGGAATTAATCTGAACTCTGGCAACAGAAAGACAAATATGGACGAGC 2460  
Db 2401 CCTGTTGAGAAAGGAATTAATCTGAACTCTGGCAACAGAAAGACAAATATGGACGAGC 2460  
Qy 2461 CAAAGAAATGCCCGTCTCTTCAAGTTAACTTAAAGGATTTCACTTCTTCCCTACCAAA 2520  
Db 2461 CAAAGAAATGCCCGTCTCTTCAAGTTAACTTAAAGGATTTCACTTCTTCCCTACCAAA 2520  
Qy 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAAAGGATTTCAAAGATTTCTTAAGGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAAAGGATTTCAAAGATTTCTTAAGGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTAGTAAACCATGCACTCTCCCTGCAGTAATTTCCGTAGTGGATTGAG 2640  
Db 2581 AAGCCCAAGGCTTAGTAAACCATGCACTCTCCCTGCAGTAATTTCCGTAGTGGATTGAG 2640  
Qy 2641 GAGGCACAGAAACCCAGTGGAGCAGTGGAGGTTAGTGCAAGATCTCAGAGATTTCAATGG 2700  
Db 2641 GAGGCACAGAAACCCAGTGGAGCAGTGGAGGTTAGTGCAAGATCTCAGAGATTTCAATGG 2700  
Qy 2701 AGGCGTTGTCTCTTTTATACCCAGCTGPACTAGCCCTTATACCTGTGMYTTTCCCAATAC 2760  
Db 2701 AGGCGTTGTCTCTTTTATACCCAGCTGPACTAGCCCTTATACCTGTGMYTTTCCCAATAC 2760  
Qy 2761 CAGAGGAAGCAGAGTGGTTTACSTCTCGGACCTTMAGGATGCCTTCTTCGATCCCTG 2820  
Db 2761 CAGAGGAAGCAGAGTGGTTTACSTCTCGGACCTTMAGGATGCCTTCTTCGATCCCTG 2820  
Qy 2821 TACATCTGACTCTCAATTTCTTGTGCTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
Db 2821 TACATCTGACTCTCAATTTCTTGTGCTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
Qy 2881 TCACCTGACTCTTTTACCCCAAGGGTTTCAAGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
Db 2881 TCACCTGACTCTTTTACCCCAAGGGTTTCAAGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
Qy 2941 TAGCCCAAGACTTGAGYCARITYMTACACTTGGACACTCTTGTCTCTCRGTAKGTGGATG 3000  
Db 2941 TAGCCCAAGACTTGAGYCARITYMTACACTTGGACACTCTTGTCTCTCRGTAKGTGGATG 3000  
Qy 3001 ATTTACTTTTGGCYGCCYRTTTCAGAAACCTTGTGCACTCAAGCCCAACCAAGCRCTTTMA 3060  
Db 3001 ATTTACTTTTGGCYGCCYRTTTCAGAAACCTTGTGCACTCAAGCCCAACCAAGCRCTTTMA 3060  
Qy 3061 ATTTCTCTCGYACCTGTGGCTTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC 3120  
Db 3061 ATTTCTCTCGYACCTGTGGCTTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC 3120  
Qy 3121 AGGTTAAATACTTAGGCTAARATTTATCCAAAGGCCACCGCCCTCAGTGAGGAAYYA 3180  
Db 3121 AGGTTAAATACTTAGGCTAARATTTATCCAAAGGCCACCGCCCTCAGTGAGGAAYYA 3180  
Qy 3181 TCCAGCTTACTTGGCTTATCTCATCYCAAAAACCTTAAAGCAACTAAGRGRTTTCCTTG 3240  
Db 3181 TCCAGCTTACTTGGCTTATCTCATCYCAAAAACCTTAAAGCAACTAAGRGRTTTCCTTG 3240  
Qy 3241 GRTTAAAYAGGTTCTGCGGAATAGGATTTCCCGAGTGTGGCRAAATAGCCAGGYCATTTA 3300  
Db 3241 GRTTAAAYAGGTTCTGCGGAATAGGATTTCCCGAGTGTGGCRAAATAGCCAGGYCATTTA 3300



3301 WATACASTAATTAAGGAACTCAGAAAGCCAATACCCATTTARTAAAGATGAYAMCTGAA 3360  
3301 WATACASTAATTAAGGAACTCAGAAAGCCAATACCCATTTARTAAAGATGAYAMCTGAA 3360  
3361 GYNRAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAAAGYCCCAGTGTTAAGYT 3420  
3361 GYNRAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAAAGYCCCAGTGTTAAGYT 3420  
3421 TGCCACRGGGCAAGACTTTTSTTAYATYRTCA CAGAAAAAACA CAGBAAVAGCTCTRGGG 3480  
3421 TGCCACRGGGCAAGACTTTTSTTAYATYRTCA CAGAAAAAACA CAGBAAVAGCTCTRGGG 3480  
3481 GTCTTTACACAGRTCCRAGGAYAGCTTGCACCCYRTGCRVACCTGASTAGGAAAYT 3540  
3481 GTCTTTACACAGRTCCRAGGAYAGCTTGCACCCYRTGCRVACCTGASTAGGAAAYT 3540  
3541 GATGTAGTGCAAGGGTTGRCVTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTIVKTA 3600  
3541 GATGTAGTGCAAGGGTTGRCVTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTIVKTA 3600  
3601 GTATCTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3601 GTATCTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3661 GTGAAYRGCAATCTCACTGCTAAAGGAGACTGTGGCTGTGCAGCAACYGTTTACTTAA 3720  
3661 GTGAAYRGCAATCTCACTGCTAAAGGAGACTGTGGCTGTGCAGCAACYGTTTACTTAA 3720  
3721 TRTCAGGCTCTAATCTGAAGGCCAGTGTCTGCRACCTGTGCACCTTGTCAAACCTTTAAC 3780  
3721 TRTCAGGCTCTAATCTGAAGGCCAGTGTCTGCRACCTGTGCACCTTGTCAAACCTTTAAC 3780  
3781 CCAGYCNATCTTCTCCAGCAATGAAGAAAGATARAAYATACTGTCAAACCAARTAAAT 3840  
3781 CCAGYCNATCTTCTCCAGCAATGAAGAAAGATARAAYATACTGTCAAACCAARTAAAT 3840  
3841 TCTCAAACCTTATCCCACTCGAGGGACCTTGTAGAGTTCCTTGTGACTGATCCYACCTT 3900  
3841 TCTCAAACCTTATCCCACTCGAGGGACCTTGTAGAGTTCCTTGTGACTGATCCYACCTT 3900  
3901 CAACTTGTATCTATGATGGAGTTCTTTGTAGAAAAAGACTTCCGAAAAGYGGGGTATGC 3960  
3901 CAACTTGTATCTATGATGGAGTTCTTTGTAGAAAAAGACTTCCGAAAAGYGGGGTATGC 3960  
3961 AGTGTCAGTGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACTAGTGCTYA 4020  
3961 AGTGTCAGTGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACTAGTGCTYA 4020  
4021 GCTRGCAAGTAATAGCCCTCACTGAGGCACTAGAAATTAGGAGAGRAAAAAAGGYAAA 4080  
4021 GCTRGCAAGTAATAGCCCTCACTGAGGCACTAGAAATTAGGAGAGRAAAAAAGGYAAA 4080  
4081 TATATATACAGACTCTRARTATGCTYACCTAGTTCNTCCATGCCCATGMRCAATATGSAR 4140  
4081 TATATATACAGACTCTRARTATGCTYACCTAGTTCNTCCATGCCCATGMRCAATATGSAR 4140  
4141 AGAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
4141 AGAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
4201 ATTATTTAYTGGWGTACAGAACTTARAGAGTGGMAGTCTTACATCTCYVGGGTATCA 4260  
4201 ATTATTTAYTGGWGTACAGAACTTARAGAGTGGMAGTCTTACATCTCYVGGGTATCA 4260  
4261 NAAAGGAAGRAAGGGAATASAGRAAYTCCCAAGCAKATATTTGAAGCHAAAGAGC 4320  
4261 NAAAGGAAGRAAGGGAATASAGRAAYTCCCAAGCAKATATTTGAAGCHAAAGAGC 4320  
4321 TGCAAGGCAAGGACCTTCCATTAGAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
4321 TGCAAGGCAAGGACCTTCCATTAGAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
4381 CTTCCGGGAAACCAAGCCCAAGTACTCAGACAGAGAAACAGAAATGGGGAACCTCACAGG 4440

4381 CTTCCGGGAAACCAAGCCCAAGTACTCAGACAGAGAAACAGAAATGGGGAACCTCACAGG 4440  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTTGCTCTGCAAC 4500  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTTGCTCTGCAAC 4500  
4501 TATCCAATGGAATATTAATAAAACCTTTCAATAAAACCTTTCACTTAGGCATCGATAGCAC 4560  
4501 TATCCAATGGAATATTAATAAAACCTTTCAATAAAACCTTTCACTTAGGCATCGATAGCAC 4560  
4561 CCATCARATGGCCAAATCATTTATTTACTGACAGGCGCTTTTCAAAACTATCAAGCARAT 4620  
4561 CCATCARATGGCCAAATCATTTATTTACTGACAGGCGCTTTTCAAAACTATCAAGCARAT 4620  
4621 AKTCAGGGCTGTGAARTGTGCCARARAATAATCCCTGCTCTYATCGCCAAAGCTCTTTC 4680  
4621 AKTCAGGGCTGTGAARTGTGCCARARAATAATCCCTGCTCTYATCGCCAAAGCTCTTTC 4680  
4681 AGGARAACAAARAACAGGCCATTTACCTGGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
4681 AGGARAACAAARAACAGGCCATTTACCTGGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
4741 CCCAAACCTCAGGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTTTACGGGTTGGGCA 4800  
4741 CCCAAACCTCAGGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTTTACGGGTTGGGCA 4800  
4801 RAGGCCCTTCCCTGTAGGACAGAAAAAGGCCAAAGAGGTAAATAAGGCACCTAGTTTCATGAA 4860  
4801 RAGGCCCTTCCCTGTAGGACAGAAAAAGGCCAAAGAGGTAAATAAGGCACCTAGTTTCATGAA 4860  
4861 ATAATTTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAG 4920  
4861 ATAATTTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAG 4920  
4921 GCCACAGTAAACCCAGGGAGTATCCAGGGCTTAGGTATAGGATATCACTTACACTGGGCC 4980  
4921 GCCACAGTAAACCCAGGGAGTATCCAGGGCTTAGGTATAGGATATCACTTACACTGGGCC 4980  
4981 TGAAGGCCACAGTCTCCTCAGGGAAGTTCGAGAAATGAATGAATAATGAATAATGAATAATGA 5040  
4981 TGAAGGCCACAGTCTCCTCAGGGAAGTTCGAGAAATGAATGAATAATGAATAATGAATAATGA 5040  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCCTGTGCTATAGCCTTAAAA 5100  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCCTGTGCTATAGCCTTAAAA 5100  
5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
5161 CCTTTCATAAACCAATGACCTTGTGTTGACCCCAAGACGCCAATTTAGTTGCAAGATCA 5220  
5161 CCTTTCATAAACCAATGACCTTGTGTTGACCCCAAGACGCCAATTTAGTTGCAAGATCA 5220  
5221 CCTCTTAGCCAAATATCAACAAAGTTCTTAACAACTTAAAGAACCTATCCCTGAGAAG 5280  
5221 CCTCTTAGCCAAATATCAACAAAGTTCTTAACAACTTAAAGAACCTATCCCTGAGAAG 5280  
5281 AGGAAAAAGAACTAATTTCCACCCWGTGACATGGTATAGTCAAGTCCCTTCCTCTAAAT 5340  
5281 AGGAAAAAGAACTAATTTCCACCCWGTGACATGGTATAGTCAAGTCCCTTCCTCTAAAT 5340  
5341 CCCCATCCCTAGATACATCTCGGGAAGGACCTTACCCAGTCAATTTATTTATACCCCACTG 5400  
5341 CCCCATCCCTAGATACATCTCGGGAAGGACCTTACCCAGTCAATTTATTTATACCCCACTG 5400  
5401 CGGTTAAAGTGGCTCGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAACTCTGGATAC 5460  
5401 CGGTTAAAGTGGCTCGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAACTCTGGATAC 5460  
5461 TGCCAAAGGAACCTGAAAAATCCAGGAGCAACCGCTAGCTATTTCTGTGAACTCTAGAGG 5520



Db	5461	TGCCAAAGGAACTGTAAAAATCCAGGAGACAAACGCTAGCTATTCTGTGAAACCTCTAGAGG	5520
Qy	5521	ATTTGGCCCTGCTCTCTCAAACAACACAGGAGGAAAGTAACATAAAATCATATAATCCCC	5580
Db	5521	ATTTGGCCCTGCTCTCTCAAACAACACAGGAGGAAAGTAACATAAAATCATATAATCCCC	5580
Qy	5581	ATGGSCTCCCTTATCATATATTTTCTCTKTAAGTTTSTTTTACCTCTTTTCACTCTCACT	5640
Db	5581	ATGGSCTCCCTTATCATATATTTTCTCTKTAAGTTTSTTTTACCTCTTTTCACTCTCACT	5640
Qy	5641	GCACCCCTCCATGCGCTGTATGACACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGA	5700
Db	5641	GCACCCCTCCATGCGCTGTATGACACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGA	5700
Qy	5701	ATGCAGCGTCCCGAAATATTGATGCCCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCC	5760
Db	5701	ATGCAGCGTCCCGAAATATTGATGCCCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCC	5760
Qy	5761	ACCTTCACTGCCCCACACCCATATATGCCCGCGCAACTGCTATCACTCTGCCACTCTTTGCATG	5820
Db	5761	ACCTTCACTGCCCCACACCCATATATGCCCGCGCAACTGCTATCACTCTGCCACTCTTTGCATG	5820
Qy	5821	CATGCAAACTCAATATTATGGAACAGGAAAAATGATTAATCTTAGTTGTCCTGAGAGCACTT	5880
Db	5821	CATGCAAACTCAATATTATGGAACAGGAAAAATGATTAATCTTAGTTGTCCTGAGAGCACTT	5880
Qy	5881	GGAGTCACTGCTCTGAGCTTACTTCAACCCAACTGGTATGCTGATGGGGGTGGAGTT	5940
Db	5881	GGAGTCACTGCTCTGAGCTTACTTCAACCCAACTGGTATGCTGATGGGGGTGGAGTT	5940
Qy	5941	CAAGATCAGGCACAGGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACAT	6000
Db	5941	CAAGATCAGGCACAGGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACAT	6000
Qy	6001	GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAAACCTCCGTAAC	6060
Db	6001	GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAAAACCTCCGTAAC	6060
Qy	6061	CATACTCGCCTGGTAAGCCTATTTTAATACACACCTCACTGGGTCCATGAGGTCTCGGCC	6120
Db	6061	CATACTCGCCTGGTAAGCCTATTTTAATACACACCTCACTGGGTCCATGAGGTCTCGGCC	6120
Qy	6121	CAAAAACCTCTAACTGTGTGATATGCTCCGCCCTGAACTTTCARGCCCATATGTTTCAATC	6180
Db	6121	CAAAAACCTCTAACTGTGTGATATGCTCCGCCCTGAACTTTCARGCCCATATGTTTCAATC	6180
Qy	6181	CCTGTACTGAAACAATGAAACAACCTTCAGCACAGAAAATAACACCACCTTCCGTTTTAGTA	6240
Db	6181	CCTGTACTGAAACAATGAAACAACCTTCAGCACAGAAAATAACACCACCTTCCGTTTTAGTA	6240
Qy	6241	GGACCTCTGTGTTTCCAAATSTGGAATAAACCATACCTTCAAACCTCACTGTGTAAAAATTT	6300
Db	6241	GGACCTCTGTGTTTCCAAATSTGGAATAAACCATACCTTCAAACCTCACTGTGTAAAAATTT	6300
Qy	6301	AGCAATACTACATACAAACAACTCCCAATGCACTAGGTGGGTAACTCTCCACACAA	6360
Db	6301	AGCAATACTACATACAAACAACTCCCAATGCACTAGGTGGGTAACTCTCCACACAA	6360
Qy	6361	ATAGTCTGCCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG	6420
Db	6361	ATAGTCTGCCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG	6420
Qy	6421	AATGGCTCTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCCCCCYATGRCATCTAC	6480
Db	6421	AATGGCTCTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCCCCCYATGRCATCTAC	6480
Qy	6481	ACTGAAACAAGATTTATACAGTATTATGTCATATCTAAGCCCCGCAACAAAAGATACCCATT	6540
Db	6481	ACTGAAACAAGATTTATACAGTATTATGTCATATCTAAGCCCCGCAACAAAAGATACCCATT	6540
Qy	6541	CTTCTTTTGTATAGGACGAGGAGTCTAGGTGCATCTAGTACTGTCATTTGGCGGTATC	6600
Db	6541	CTTCTTTTGTATAGGACGAGGAGTCTAGGTGCATCTAGTACTGTCATTTGGCGGTATC	6600

Qy	6601	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATCGAACGG	6660
Db	6601	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATCGAACGG	6660
Qy	6661	GTCCGCGACCTCCCTGTCTACCTTGCAGAGTCAACTAACTCCCTAGCAGCAGTAGTCCTT	6720
Db	6661	GTCCGCGACTCCCTGGTCTACCTTGCAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCCTT	6720
Qy	6721	CRAAATCGAAGAGCTTTAGACTGCTTAAACCGCTGARAGAGGGGGAACCTGTTATTTTAA	6780
Db	6721	CRAAATCGAAGAGCTTTAGACTTGTCTTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA	6780
Qy	6781	GGGGAAGAATGCTGTTATTTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA	6840
Db	6781	GGGGAAGAATGCTGTTATTTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTRAAGAA	6840
Qy	6841	ATTCSAGATCGAATAACAAGTAKAGCAGARGAGCTTCGAAACACTGGACCTCGGGGCTC	6900
Db	6841	ATTCSAGATCGAATAACAAGTAKAGCAGARGAGCTTCGAAACACTGGACCTCGGGGCTC	6900
Qy	6901	CTCAGCCBRATGGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG	6960
Db	6901	CTCAGCCBRATGGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG	6960
Qy	6961	CTACTCTCTTTTGGACCCCTGTATCTTTTACCTCTCTTGTAACTTTGTCTCTCCAGAATC	7020
Db	6961	CTACTCTCTTTTGGACCCCTGTATCTTTTACCTCTCTTGTAACTTTGTCTCTCCAGNATC	7020
Qy	7021	GAAGCTGTAAACTCAAAATGGAGCCCAAGATGCAGTCCAAGACTAAGACTCTACCGCAGA	7080
Db	7021	GAAGCTGTAAACTCAAAATGGAGCCCAAGATGCAGTCCAAGACTAAGACTCTACCGCAGA	7080
Qy	7081	CCCTTGGACCGGCTGYTAGCCACAGATCTGATGTTTAAATGACATCAAAGGACCCCTCTCT	7140
Db	7081	CCCTTGGACCGGCTGYTAGCCACAGATCTGATGTTTAAATGACATCAAAGGACCCCTCTCT	7140
Qy	7141	GAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTTACAGCAGGAAGCAGTTAGAGC	7200
Db	7141	GAGGAATCTCAGCTGCACAACTCTACTACGCCCCNATTTACAGCAGGAAGCAGTTAGAGC	7200
Qy	7201	GGTSGTCGGCCAACTCCCAACAGCACTTAPAGTTTTCTGTGTAGATGGGGACTGAGA	7260
Db	7201	GGTSGTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTGTAGATGGGGACTGAGA	7260
Qy	7261	GACAGNACTAGCTGGATTTCTTAGCTGATYTAGAATCCYTAGCCTAGSTGGGAGGTG	7320
Db	7261	GACAGGACTAGCTGGATTTCTTAGCTGATYTAGAATCCYTAGAAGTCCYTAGACCTAGSTGGGAGGTG	7320
Qy	7321	ACCACATCCACCTTTAAACACGGGGCTTGCAACTTTAGYTCAACACCTGACCAATCAGAGAG	7380
Db	7321	ACCACATCCACCTTTAAACACGGGGCTTGCAACTTTAGYTCAACACCTGACCAATCAGAGAG	7380
Qy	7381	CTCACTAAAAATGCTAATTAGGCGAAGAAGAGAGGTAAAGAAATAGCCAAATCATYATTGC	7440
Db	7381	CTCACTAAAAATGCTAATTAGGCGAAGAAGAGAGGTAAAGAAATAGCCAAATCATYATTGC	7440
Qy	7441	MTGAGAGCAGCAGGAGGGAACAATGATCCGGATATAAACCCAAAGTTTTCGAGCCGGCAA	7500
Db	7441	MTGAGAGCAGCAGGAGGGAACAATGATCCGGATATAAACCCAAAGTTTTCGAGCCGGCAA	7500
Qy	7501	CGGCACCCCTTTGGTCCCTCCCTTCTATGGGAGCTCTGTTTTTCATGCTATTTCAC	7560
Db	7501	CGGCACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCAC	7560
Qy	7561	TCATTAAATCTTGCARCTGCR 7582	
Db	7561	TCATTAAATCTTGCARCTGCR 7582	
RESULT 4			
AX027480			
LOCUS			
AX027480		7582 bp	DNA linear
			PAT 16-SEP-2000

DEFINITION Sequence 30 from Patent FR2788784.  
ACCESSION AX027480  
VERSION AX027480.1 GI:10188444  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Mallet, F., Voisset, C. and Paranhos, B.G.  
JOURNAL Patent: FR 2788784-A 30 28-JUL-2000;  
BIO MERIEUX (FR)  
FEATURES  
Location/Qualifiers  
1..7582  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 7489.6; DB 2; Length 7582;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAAGCAGAGCCGCCCTTTGGGTCC 60  
DB 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAAGCAGAGCCGCCCTTTGGGTCC 60  
QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCAACTGCA 120  
DB 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCAACTGCA 120  
QY 121 CTTCTTGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180  
DB 121 CTTCTTGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180  
QY 181 TGTTCGCCACCCAGCANACCTGCGCTGACTCCCATCCCTCTGGATCTCGAGGCTGTC 240  
DB 181 TGTTCGCCACCCAGCANACCTGCGCTGACTCCCATCCCTCTGGATCTCGAGGCTGTC 240  
QY 241 CGCTGTGCTCTGATCCAGCGARGCGCCATTCGCGCTCCCAATGGGCTTAAAGGCTGCG 300  
DB 241 CGCTGTGCTCTGATCCAGCGARGCGCCATTCGCGCTCCCAATGGGCTTAAAGGCTGCG 300  
QY 301 CATTTGNTCTGCAAGCTAAGTGCCTGGGTTTGTCTTAATGAGCTGAACTANTCACT 360  
DB 301 CATTTGNTCTGCAAGCTAAGTGCCTGGGTTTGTCTTAATGAGCTGAACTANTCACT 360  
QY 361 GGGTTCCATGGTCTCTTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTACCACA 420  
DB 361 GGGTTCCATGGTCTCTTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTACCACA 420  
QY 421 TGGCCCAAGATTCATTCTTGGAAATCCGTGAGGSCAACGAATCCAGGTCAGAGATAC 480  
DB 421 TGGCCCAAGATTCATTCTTGGAAATCCGTGAGGSCAACGAATCCAGGTCAGAGATAC 480  
QY 481 GARGCTTGCCACATCTTGGAAAGCGGCTGCTACRCTTTGGAAGTGGTTTCCACCACATC 540  
DB 481 GARGCTTGCCACATCTTGGAAAGCGGCTGCTACRCTTTGGAAGTGGTTTCCACCACATC 540  
QY 541 TTGGGAGCTCTGTGAGCAGGACCCCGGTRACATTTTGGGRCACAMSRACGACATCC 600  
DB 541 TTGGGAGCTCTGTGAGCAGGACCCCGGTRACATTTTGGGRCACAMSRACGACATCC 600  
QY 601 MAAGTATGGGAAACCTTCCCGCAGACAGAAAACGCCCTTAAGCTATTCTGGARANT 660  
DB 601 MAAGTATGGGAAACCTTCCCGCAGACAGAAAACGCCCTTAAGCTATTCTGGARANT 660  
QY 661 TGGGAMCAATTTGACCTCTCAGACACTAAGAAAGAAACGACTTATTTCTTCTGCACTGCC 720  
DB 661 TGGGAMCAATTTGACCTCTCAGACACTAAGAAAGAAACGACTTATTTCTTCTGCACTGCC 720  
QY 721 GCCTGGCACTCTCGAGGGAAGTATAAATTAACCACTTTACAGCTAGACTCTTTTGG 780

DB 721 GCCTGGCACTCTCGAGGGAAGTATAAATTAACCACTTTACAGCTAGACTCTTTTGG 780  
QY 781 TAGAAAAGGCAAAATGAGTGAAGTGCATAAAGTACAAAATTTCTTTTTCATTAGAGACAA 840  
DB 781 TAGAAAAGGCAAAATGAGTGAAGTGCATAAAGTACAAAATTTCTTTTTCATTAGAGACAA 840  
QY 841 CTCACAATTATGTAAAAAAGTGTGATTTATGCCCTCAGGAAGCTTTCAGAGTCTACCTCC 900  
DB 841 CTCACAATTATGTAAAAAAGTGTGATTTATGCCCTCAGGAAGCTTTCAGAGTCTACCTCC 900  
QY 901 CTATCCACGATCCCGGACTCTTCCCCAMYTAAAGGACCCCTTCAACCAATAGG 960  
DB 901 CTATCCACGATCCCGGACTCTTCCCCAMYTAAAGGACCCCTTCAACCAATAGG 960  
QY 961 TCCAAAAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCATATTTCCCAAT 1020  
DB 961 TCCAAAAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCATATTTCCCAAT 1020  
QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCACGACAGTGCATGTGCT 1080  
DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCACGACAGTGCATGTGCT 1080  
QY 1081 TTTTCTCCGAGCTTAAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAATCT 1140  
DB 1081 TTTTCTCCGAGCTTAAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAATCT 1140  
QY 1141 GATGGCTATATTTGRTCTTTTCAAGGGTTAGGACAATTTCTTGATCTGACATGGAGAT 1200  
DB 1141 GATGGCTATATTTGRTCTTTTCAAGGGTTAGGACAATTTCTTGATCTGACATGGAGAT 1200  
QY 1201 ATATATGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCACCACTAATG 1260  
DB 1201 ATATATGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCACCACTAATG 1260  
QY 1261 AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGATGATGATGACAA 1320  
DB 1261 AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGATGATGATGACAA 1320  
QY 1321 GAAGGAAGANAATGATTTCCCAAGGCGCAGCAGGAGTCCAGCTTCCAGCTTACCTCAT 1380  
DB 1321 GAAGGAAGANAATGATTTCCCAAGGCGCAGCAGGAGTCCAGCTTACCTCAT 1380  
QY 1381 GGGACACAGAAATCAGTAACTAGTGGAGATTTGGTCTGACAGACATTTGCTAACTGTG 1440  
DB 1381 GGGACACAGAAATCAGTAACTAGTGGAGATTTGGTCTGACAGACATTTGCTAACTGTG 1440  
QY 1441 TASAAGGACTAAGGAAAACCTASGAAGAAATCTAYGAAATTAATCAATGATGTCACCA 1500  
DB 1441 TASAAGGACTAAGGAAAACCTASGAAGAAATCTAYGAAATTAATCAATGATGTCACCA 1500  
QY 1501 ACACAGGGAAGGGAAGAAAATCCTACTGCTTCTGAGAGACTTAAGGAGGCAATGAG 1560  
DB 1501 ACACAGGGAAGGGAAGAAAATCCTACTGCTTCTGAGAGACTTAAGGAGGCAATGAG 1560  
QY 1561 GAAGGCTGCTCTCTGTCACTGATCTTTCTGAAGGCCAACTAACTTTAAAGCGTAAGT 1620  
DB 1561 GAAGGCTGCTCTCTGTCACTGATCTTTCTGAAGGCCAACTAACTTTAAAGCGTAAGT 1620  
QY 1621 TATCACTCAGTGCAGACATTTAGAAAAAATTTCAAAAGTCTGCCGTAGGCCCGGAG 1680  
DB 1621 TATCACTCAGTGCAGACATTTAGAAAAAATTTCAAAAGTCTGCCGTAGGCCCGGAG 1680  
QY 1681 CAAAACCTTAGAACCTTATGACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACT 1740  
DB 1681 CAAAACCTTAGAACCTTATGACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACT 1740  
QY 1741 GAGCAGGCGGAACAGACAAACCGGATTTAAAAAAGGCCCGCTTTAGTCAATGACCT 1800  
DB 1741 GAGCAGGCGGAACAGACAAACCGGATTTAAAAAAGGCCCGCTTTAGTCAATGACCT 1800  
QY 1801 CAGGCAAGTGGACTTTTGGAGGCTCTGGAAAAGGGAAGTCTGGGCAAAATTTGAATGCT 1860

Db 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAGGAAAAAGCTGGGCAAAATGAAATGCCTAA 1860  
Qy 1861 TAGGCGTCTGCTTCCAGTGGGCTCTCAAGGACACTTTTAAAAAGATTTGTCCAAAGTAGAAG 1920  
Db 1861 TAGGCGTCTGCTTCCAGTGGGCTCTCAAGGACACTTTTAAAAAGATTTGTCCAAAGTAGAAG 1920  
Qy 1921 TAAGCGCGCCCTTCCTCCATGCCCTTATTTCAAGGGATCACTGGAAGGCCCACTGCC 1980  
Db 1921 TAAGCGCGCCCTTCCTCCATGCCCTTATTTCAAGGGATCACTGGAAGGCCCACTGCC 1980  
Qy 1981 CAGGGACAAGGCTCTTTTGAAGTCAAGAGCCCACTAACCAAGATGATCCAGCAGCAGACTG 2040  
Db 1981 CAGGGACAAGGCTCTTTTGAAGTCAAGAGCCCACTAACCAAGATGATCCAGCAGCAGACTG 2040  
Qy 2041 AGGTCCTGGGGCAAGGCCCATCCATGCCATCACCTCCACAGAGCCCTGGGTATGCTT 2100  
Db 2041 AGGTCCTGGGGCAAGGCCCATCCATGCCATCACCTCCACAGAGCCCTGGGTATGCTT 2100  
Qy 2101 GACCAATTGAGGCGCAGGAAGTGTCTCTCGACACTGGTGGGTCTTCTTAGTCTTACT 2160  
Db 2101 GACCAATTGAGGCGCAGGAAGTGTCTCTCGACACTGGTGGGTCTTCTTAGTCTTACT 2160  
Qy 2161 CTTCTGTCGCGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACG 2220  
Db 2161 CTTCTGTCGCGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACG 2220  
Qy 2221 GGCAGTCACTAGATCTTTTCCAGCCCACTAAGTATGAATGGGAGCTTTATCTTT 2280  
Db 2221 GGCAGTCACTAGATCTTTTCCAGCCCACTAAGTATGAATGGGAGCTTTATCTTT 2280  
Qy 2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTTGAGGAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTTGAGGAGAGACATTTCTA 2340  
Qy 2341 GCAAAAGCAGGGCCCACTATACCTGAACATAGAGAAGAAACCCGTTTGTGTNCC 2400  
Db 2341 GCAAAAGCAGGGCCCACTATACCTGAACATAGAGAAGAAACCCGTTTGTGTNCC 2400  
Qy 2401 CTTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGCAATATGGACGAGC 2460  
Db 2401 CTTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGCAATATGGACGAGC 2460  
Qy 2461 CAAGAATGCCGCTCTGTTCAAGTTAACTAAGGATTCACCTTCCCTACCATA 2520  
Db 2461 CAAGAATGCCGCTCTGTTCAAGTTAACTAAGGATTCACCTTCCCTACCATA 2520  
Qy 2521 GGCAGTACCCCTCAGACCAAGGCCCAAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCAAGGCCCAAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTAGTAAACCATGCACTAATCCCTGCAAGTAAATCCGTTAGGATGAG 2640  
Db 2581 AAGCCCAAGGCTTAGTAAACCATGCACTAATCCCTGCAAGTAAATCCGTTAGGATGAG 2640  
Qy 2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGTTAGTGCAAGATCTCAGGATTAATCAATGG 2700  
Db 2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGTTAGTGCAAGATCTCAGGATTAATCAATGG 2700  
Qy 2701 AGGCGGTTGCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCGCCAAATAC 2760  
Db 2701 AGGCGGTTGCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCGCCAAATAC 2760  
Qy 2761 CAGAGGAACAGAGTGGTTTACASTCTCGACCTTMAAGATGCCCTTCTCGATCCCTG 2820  
Db 2761 CAGAGGAACAGAGTGGTTTACASTCTCGACCTTMAAGATGCCCTTCTCGATCCCTG 2820  
Qy 2821 TACATCTCACTCTCAATCTTTGTTGCTTTGAGATACTTCAACCCCATCTCAAC 2880  
Db 2821 TACATCTCACTCTCAATCTTTGTTGCTTTGAGATACTTCAACCCCATCTCAAC 2880  
Qy 2881 TCACCTGGACTTTTACCCCAAGGGTTCCAGGGATAGYCCCATCTATTATGGCCAGGAT 2940  
Db 2881 TCACCTGGACTTTTACCCCAAGGGTTCCAGGGATAGYCCCATCTATTATGGCCAGGAT 2940

Qy 2941 TAGCCCAAGACTTGAGYCAPTYMTACATCTGGACACTCTTGTCTCTCRGTAKGTGGATG 3000  
Db 2941 TAGCCCAAGACTTGAGYCAPTYMTACATCTGGACACTCTTGTCTCTCRGTAKGTGGATG 3000  
Qy 3001 ATTTACTTTTTRGCGCCYRTTCAGAAACCTTGTGCATCAAGGCCACCCAAAGCRCTCTTMA 3060  
Db 3001 ATTTACTTTTTRGCGCCYRTTCAGAAACCTTGTGCATCAAGGCCACCCAAAGCRCTCTTMA 3060  
Qy 3061 ATTTCTCGYACCTGTGGCTACAWGGTTTCCAAACCSARAGCTCARCTCTGCTCACAGC 3120  
Db 3061 ATTTCTCGYACCTGTGGCTACAWGGTTTCCAAACCSARAGCTCARCTCTGCTCACAGC 3120  
Qy 3121 AGGTTAAATACCTTAGGCTAARATTTATCCAAAGGCCACCCAGGCCCTCAGTGAGGAAYRA 3180  
Db 3121 AGGTTAAATACCTTAGGCTAARATTTATCCAAAGGCCACCCAGGCCCTCAGTGAGGAAYRA 3180  
Qy 3181 TCCAGCCTATCTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRGRTTCTTTG 3240  
Db 3181 TCCAGCCTATCTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRGRTTCTTTG 3240  
Qy 3241 GCRTAAAYAGGYTTCTGCCGAATWATGGATTCCCAGGTTWGCRAAATAGCCAGGYCATTA 3300  
Db 3241 GCRTAAAYAGGYTTCTGCCGAATWATGGATTCCCAGGTTWGCRAAATAGCCAGGYCATTA 3300  
Qy 3301 WATACASTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTTARTAAGATGGAYAMCTGAA 3360  
Db 3301 WATACASTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTTARTAAGATGGAYAMCTGAA 3360  
Qy 3361 GYMAAGTGGCTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCAGTGTTAAGYT 3420  
Db 3361 GYMAAGTGGCTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCAGTGTTAAGYT 3420  
Qy 3421 TGCCAACRGGCAAGACTTTTATAYRTACAGAAAAAACAAGRAAYAGCTCTRGA 3480  
Db 3421 TGCCAACRGGCAAGACTTTTATAYRTACAGAAAAAACAAGRAAYAGCTCTRGA 3480  
Qy 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTCAACCCYRTGGCRYACCTGASTAAGGAAYT 3540  
Db 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTCAACCCYRTGGCRYACCTGASTAAGGAAYT 3540  
Qy 3541 GATGTAGTGGCAAGGGTGTGCTTATTTAYGGGTAGTGGCTGAGTACAGTYKTA 3600  
Db 3541 GATGTAGTGGCAAGGGTGTGCTTATTTAYGGGTAGTGGCTGAGTACAGTYKTA 3600  
Qy 3601 GTATCTGAAGCAGTTTAAATTAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Db 3601 GTATCTGAAGCAGTTTAAATTAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Qy 3661 GTGAAYRCATACCTCAGTCTTAAAGGACCTTGTGGCTGTGACAAACCTTACTTAA 3720  
Db 3661 GTGAAYRCATACCTCAGTCTTAAAGGACCTTGTGGCTGTGACAAACCTTACTTAA 3720  
Qy 3721 TRTCAAGGCTTATTTGAAGGGCAGTGTGCTGCTGCTGCTGCTGCTGCTTAAAC 3780  
Db 3721 TRTCAAGGCTTATTTGAAGGGCAGTGTGCTGCTGCTGCTGCTGCTTAAAC 3780  
Qy 3781 CCAGYCNCACTTTCTCCAGACAAATGAAGAAAGATARAAYATACTGTCAAACAAATTAAT 3840  
Db 3781 CCAGYCNCACTTTCTCCAGACAAATGAAGAAAGATARAAYATACTGTCAAACAAATTAAT 3840  
Qy 3841 TCTCAAACTATGCCACTCGAGGGGACCTTTTAGAGTTCCCTTACCTGATCCYGACCTT 3900  
Db 3841 TCTCAAACTATGCCACTCGAGGGGACCTTTTAGAGTTCCCTTACCTGATCCYGACCTT 3900  
Qy 3901 CAACCTGTATCTAGTGAAGTTCTTTGTAGAAAAAGGACTTCAAAAAAGYGGGTATGTC 3960  
Db 3901 CAACCTGTATCTAGTGAAGTTCTTTGTAGAAAAAGGACTTCAAAAAAGYGGGTATGTC 3960  
Qy 3961 AGTGTGCTAGTGAATTAATGGAATTAATGGAAGTAATCCCTCACTCCAGGAACCTAGTCTTAA 4020  
Db 3961 AGTGTGCTAGTGAATTAATGGAATTAATGGAAGTAATCCCTCACTCCAGGAACCTAGTCTTAA 4020

QY 4021 GCTRGCAACTAATAGCCYTCAYTKGGCAGCTAGTAATTAGGAGAAAGGAAAGGGYAAA 4080  
DB 4021 GCTRGCAAGACTAATAGCCYTCAYTKGGCAGCTAGTAATTAGGAGAAAGGAAAGGGYAAA 4080  
QY 4081 TATATATACAGACTCTTARTATGCTYACCTAGTCTCCTATCCATGAGCAATATGSGAR 4140  
DB 4081 TATATATACAGACTCTTARTATGCTYACCTAGTCTCCTATCCATGAGCAATATGSGAR 4140  
QY 4141 AGAAGGGAATCTCTAATCTTCYGNRGAACACCTATCAMAATCAGGAAGCCATTAGGAR 4200  
DB 4141 AGAAGGGAATCTCTAATCTTCYGNRGAACACCTATCAMAATCAGGAAGCCATTAGGAR 4200  
QY 4201 ATTATTAATGCGGTACAGAACTTARAGAGTGGAGTCTTACACTGCGYGGGTATCA 4260  
DB 4201 ATTATTAATGCGGTACAGAACTTARAGAGTGGAGTCTTACACTGCGYGGGTATCA 4260  
QY 4261 NAAAGGAAAGAAAGGGAATASAAAGRAAYTGCACAGCAKATATTGAAGCMMAAAGAGC 4320  
DB 4261 NAAAGGAAAGAAAGGGAATASAAAGRAAYTGCACAGCAKATATTGAAGCMMAAAGAGC 4320  
QY 4321 TGCAAGGCAGGACCTCCATAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTATCC 4380  
DB 4321 TGCAAGGCAGGACCTCCATAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTATCC 4380  
QY 4381 CTTCCGGGAAACCAAGCCAGTACTCAGCAGAGAAACAGAAATGSGGAACCTCAGAGG 4440  
DB 4381 CTTCCGGGAAACCAAGCCAGTACTCAGCAGAGAAACAGAAATGSGGAACCTCAGAGG 4440  
QY 4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTGCTGCAAC 4500  
DB 4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTGCTGCAAC 4500  
QY 4501 TATCCAAATGAAATTAATTAACCCCTTCAACAACTTTCACTTAGGCATCGATAGCAC 4560  
DB 4501 TATCCAAATGAAATTAATTAACCCCTTCAACAACTTTCACTTAGGCATCGATAGCAC 4560  
QY 4561 CCATCARATGCGCAAACTATTAATTAAGTGGACAGGCTTTTCAAACTATCAAGCARAT 4620  
DB 4561 CCATCARATGCGCAAACTATTAATTAAGTGGACAGGCTTTTCAAACTATCAAGCARAT 4620  
QY 4621 AKTCAGGCGCTGTGAATGTGCCARARAATAATCCCTGCTCTATCGCCAAAGCTCTTC 4680  
DB 4621 AKTCAGGCGCTGTGAATGTGCCARARAATAATCCCTGCTCTATCGCCAAAGCTCTTC 4680  
QY 4681 AGARAACAARAACAGGCAATTAACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
DB 4681 AGARAACAARAACAGGCAATTAACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
QY 4741 CCCAAACCTCAGGATTTAGTATCTACTAGTCTGGTATATATCTTTCACGGGTTGGGCA 4800  
DB 4741 CCCAAACCTCAGGATTTAGTATCTACTAGTCTGGTATATATCTTTCACGGGTTGGGCA 4800  
QY 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATGAAGCACTAGTTTCATGAA 4860  
DB 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATGAAGCACTAGTTTCATGAA 4860  
QY 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGTTCCAG 4920  
DB 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGTTCCAG 4920  
QY 4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATAGATATCACTTTACACTCGGCC 4980  
DB 4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATAGATATCACTTTACACTCGGCC 4980  
QY 4981 TGAAGCCACAGTCTCAGGGAAGTTCGAAATGAATGAAYACTCAAGGACATCTA 5040  
DB 4981 TGAAGCCACAGTCTCAGGGAAGTTCGAAATGAATGAAYACTCAAGGACATCTA 5040  
QY 5041 AAAAAGCAACCCAGGAAACCCACCTCACAATGCGCTGCTGTTGCTATAGCCTTAAAA 5100  
DB 5041 AAAAAGCAACCCAGGAAACCCACCTCACAATGCGCTGCTGTTGCTATAGCCTTAAAA 5100  
QY 5101 AGAATCTGCAACTTTTCCCAAAAGCAGGACTTTAGCCCATACGAAATGCTGTATGGAAGG 5160

DB 5101 AGAATCTGCAACTTTTCCCAAAAGCAGGACTTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
QY 5161 CCCTTCATTAACCAATGACCTTGTGCTTGACCCCAAGCAGCAACTTGTAGTTGCAGACATCA 5220  
DB 5161 CCCTTCATTAACCAATGACCTTGTGCTTGACCCCAAGCAGCAACTTGTAGTTGCAGACATCA 5220  
QY 5221 CTTCTCTAGCCAAATATCAACAAAGTCTTAAACATATTAAGGAACTTATCCCTGAGAAG 5280  
DB 5221 CTTCTCTAGCCAAATATCAACAAAGTCTTAAACATATTAAGGAACTTATCCCTGAGAAG 5280  
QY 5281 AGGGAAGAAAGTATTTCCACCCWGTGACATGTTAGTCAAGTCCCTCYCTCTAATT 5340  
DB 5281 AGGGAAGAAAGTATTTCCACCCWGTGACATGTTAGTCAAGTCCCTCYCTCTAATT 5340  
QY 5341 CCCCATCCCTAGATACATCTCTGGGAAAGGACCTTACCCAGTCACTTTTATTTACCCCAACTG 5400  
DB 5341 CCCCATCCCTAGATACATCTCTGGGAAAGGACCTTACCCAGTCACTTTTATTTACCCCAACTG 5400  
QY 5401 CGTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCGGATAC 5460  
DB 5401 CGTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCGGATAC 5460  
QY 5461 TGCCAAAGGAACCTGAAATCCAGGAGACAACTAGTCTTCTGTGACCTCTAGAGG 5520  
DB 5461 TGCCAAAGGAACCTGAAATCCAGGAGACAACTAGTCTTCTGTGACCTCTAGAGG 5520  
QY 5521 ATTTGCGCTGCTCTTCAAAACCAACACAGGAGGAAGTAACTAAATATAAATCCCCC 5580  
DB 5521 ATTTGCGCTGCTCTTCAAAACCAACACAGGAGGAAGTAACTAAATATAAATCCCCC 5580  
QY 5581 ATGSGCTCCCTTATCATATTTTCTKTASTGTTSTTTTACCTSTTTCACTCTCACT 5640  
DB 5581 ATGSGCTCCCTTATCATATTTTCTKTASTGTTSTTTTACCTSTTTCACTCTCACT 5640  
QY 5641 GCACCCCTCCATGCGCTGTATGACAGTACTCCCTTACCCCTTATGAGGATTTCTATGAGA 5700  
DB 5641 GCACCCCTCCATGCGCTGTATGACAGTACTCCCTTACCCCTTATGAGGATTTCTATGAGA 5700  
QY 5701 ATGACAGCTCCCGAAATATTTGATGCCCATCTATAGAGTCTTTTSTAAGGGAACCCCC 5760  
DB 5701 ATGACAGCTCCCGAAATATTTGATGCCCATCTATAGAGTCTTTTSTAAGGGAACCCCC 5760  
QY 5761 ACCTTCACTGCCACACCCCATATGCGCCCACTGCTATCACTCTGCGCACTCTTTGCAATG 5820  
DB 5761 ACCTTCACTGCCACACCCCATATGCGCCCACTGCTATCACTCTGCGCACTCTTTGCAATG 5820  
QY 5821 CATGCAAACTACTATTTTGGACAGGAAAAATGATTAATCTCTAGTTGCTCTGAGGACTT 5880  
DB 5821 CATGCAAACTACTATTTTGGACAGGAAAAATGATTAATCTCTAGTTGCTCTGAGGACTT 5880  
QY 5881 GGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGAGGCTGAGTT 5940  
DB 5881 GGAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGAGGCTGAGTT 5940  
QY 5941 CAAGATCAGGCAAGGAAAAACATGTAAAGAAAGTAACTCTCCAACTCAACCCGGGTACAT 6000  
DB 5941 CAAGATCAGGCAAGGAAAAACATGTAAAGAAAGTAACTCTCCAACTCAACCCGGGTACAT 6000  
QY 6001 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 6060  
DB 6001 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 6060  
QY 6061 CATACTCGCTGGTAAAGCTTAAATACCACTCAGCTGGGTCCATGAGTCTCGGCC 6120  
DB 6061 CATACTCGCTGGTAAAGCTTAAATACCACTCAGCTGGGTCCATGAGTCTCGGCC 6120  
QY 6121 CAAAACCTTAACTGTTGGATATGCGCTCCCTGAACTTCARGGCATATGTTTCAATC 6180  
DB 6121 CAAAACCTTAACTGTTGGATATGCGCTCCCTGAACTTCARGGCATATGTTTCAATC 6180  
QY 6181 CTGTGACTGAAACATTTGGAACAACTTACGACAGAAATAAAACCACTTCCGTTTTAGTA 6240

Db 6181 CCTGTACTCTGAACAATGGAACAACTTCAGCACAGAAATAAACACACACTTCCTCGTTTTAGTA 6240  
Qy 6241 GGACCTCTTGTTTCCAAATSTGAANAATACCCATACCTCAAACCTCACCTGTGTAAATTT 6300  
Db 6241 GGACCTCTTGTTTCCAAATSTGAANAATACCCATACCTCAAACCTCACCTGTGTAAATTT 6300  
Qy 6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGTAACTCCTCCACACAA 6360  
Db 6301 AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGTAACTCCTCCACACAA 6360  
Qy 6361 ATAGTCTGCCTACCTTCAGGAATATTTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTG 6420  
Db 6361 ATAGTCTGCCTACCTTCAGGAATATTTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTG 6420  
Qy 6421 AATGCTCTTCAGATCTATGCTTCCTCTCATTTCTTAGTGCCCCCATGRCATCTAC 6480  
Db 6421 AATGCTCTTCAGATCTATGCTTCCTCTCATTTCTTAGTGCCCCCATGRCATCTAC 6480  
Qy 6481 ACTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
Db 6481 ACTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
Qy 6541 CTTCTTTTGTATAGGACGAGTGCTAGTGCACTAGTAGTACCTGGCAATGGCGGTATC 6600  
Db 6541 CTTCTTTTGTATAGGACGAGTGCTAGTGCACTAGTAGTACCTGGCAATGGCGGTATC 6600  
Qy 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Db 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Qy 6661 GTCCGGCAGCTCCCTGCTGCTCAAGATCAACTTAACTCCCTAGCAGCAGTACCTT 6720  
Db 6661 GTCCGGCAGCTCCCTGCTGCTCAAGATCAACTTAACTCCCTAGCAGCAGTACCTT 6720  
Qy 6721 CAAATCGAAGAGCTTTAGACTGCTAAACCGCTGARAGAGGGGAACTGTTTATTTTA 6780  
Db 6721 CAAATCGAAGAGCTTTAGACTGCTAAACCGCTGARAGAGGGGAACTGTTTATTTTA 6780  
Qy 6781 GGGGAAGATGCTGTTATTTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
Db 6781 GGGGAAGATGCTGTTATTTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
Qy 6841 ATTCSAGATCGAATCAACGTAAGCAGAGAGCTTCGAAACACTGGACCCCTGGGCGCTC 6900  
Db 6841 ATTCSAGATCGAATCAACGTAAGCAGAGAGCTTCGAAACACTGGACCCCTGGGCGCTC 6900  
Qy 6901 CTCAGCCRAATGATCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Db 6901 CTCAGCCRAATGATCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Qy 6961 CTACTCCTCTTGGACCCCTGATCTTTTACCTCCTTGTAACTTTGTCTCTTCCAGAAATC 7020  
Db 6961 CTACTCCTCTTGGACCCCTGATCTTTTACCTCCTTGTAACTTTGTCTCTTCCAGAAATC 7020  
Qy 7021 GAAGCTGTAAACTCAAAATGAGCCCAAGATGCAAGTCAAGACTAAAGACTTACCGCAGA 7080  
Db 7021 GAAGCTGTAAACTCAAAATGAGCCCAAGATGCAAGTCAAGACTTACCGCAGA 7080  
Qy 7081 CCCCTGGACCGGCTGYTAGCCACGATCTGATGTTAAATGACATCAAAAGGACCCCTCCT 7140  
Db 7081 CCCCTGGACCGGCTGYTAGCCACGATCTGATGTTAAATGACATCAAAAGGACCCCTCCT 7140  
Qy 7141 GAGGAAATCTAGCTGCAACACTCTATACGCCCAATTCAGCAGGAGCAAGTTAGAGC 7200  
Db 7141 GAGGAAATCTAGCTGCAACACTCTATACGCCCAATTCAGCAGGAGCAAGTTAGAGC 7200  
Qy 7201 GGTGTGGCCCAACTCCCAACAGCAGTCTAGGTTTCTGTTGAGATGGGGAGCTGAGA 7260  
Db 7201 GGTGTGGCCCAACTCCCAACAGCAGTCTAGGTTTCTGTTGAGATGGGGAGCTGAGA 7260  
Qy 7261 GACGAGCTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTAGTGGGAAGGTG 7320  
Db 7261 GACGAGCTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTAGTGGGAAGGTG 7320

Qy 7321 ACCCATCCACTTTAAACACGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380  
Db 7321 ACCCATCCACTTTAAACACGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380  
Qy 7381 CTCCTAAATATGCTAAATTTAGGCAAGAGACAGAGGTAAAGAAATAGCCAAATCATYTTATTCG 7440  
Db 7381 CTCCTAAATATGCTAAATTTAGGCAAGAGACAGAGGTAAAGAAATAGCCAAATCATYTTATTCG 7440  
Qy 7441 MTGAGAGCACAGCAGGAGGGAACAATGATCGGGATATAAACCAGTYTTCGAGCGCGCAA 7500  
Db 7441 MTGAGAGCACAGCAGGAGGGAACAATGATCGGGATATAAACCAGTYTTCGAGCGCGCAA 7500  
Qy 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Db 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Qy 7561 TCTATTAATCTTGCARCTGCR 7582  
Db 7561 TCTATTAATCTTGCARCTGCR 7582

RESULT 5  
AY101582  
LOCUS  
DEFINITION  
Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus, allele A, complete sequence.  
ACCESSION  
AY101582  
VERSION  
AY101582.1 GI:37544399  
KEYWORDS  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
REFERENCE  
1 (bases 1 to 10222)  
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.  
The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology  
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)  
14757826  
2 (bases 1 to 10222)  
Mallet, F., Bouton, O. and Oriol, G.  
Direct Submission  
Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France  
JOURNAL  
Location/Qualifiers  
FEATURES  
source  
1..10222  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="21"  
/isolation\_source="PBMC of individual 21"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q21-q22"  
/sex="female"  
/cell\_type="PBMC"  
/note="isolated by PCR; ERVWE1 locus, allele A endogenous\_virus: HERV-W"  
1..780  
join(248..811,2887..9813)  
/note="putative mRNA transcript 1; gag-pol pseudogene"  
join(248..811,7585..9813)  
/product="envelope glycoprotein"  
/note="putative mRNA transcript 2"  
join(248..811,9248..9813)  
/note="putative mRNA transcript 3; no coding region"  
7812..9428  
/note="syncytin"  
/codon\_start=1  
/product="envelope glycoprotein"

Query Match	86.1%	Score 6448	DB 5	Length 10222
Best Local Similarity	95.7%	Pred. No. 0		
Matches 6681	Conservative 195	Mismatches 50	Indels 53	Gaps 27
Qy	606	GATGGGAAACGTTCCCGCAAGACAAAAAAGCGCCCTTAAGACGTATTTCTGGAATAATTGGGA	665	
Db	2886	GATGGGAAACGTTCCCGCAAGACAAAAAAGCGCCCTTAAGACGTATTTCTGGAATAATTGGGA	2945	
Qy	666	MCAATTTGACCCCTCAGACACTTAAGAAAGAAAACGACTTATATTTCTTCGTGAGTGC CGCCTG	725	
Db	2946	CCAATTTGACCCCTCAGACACTTAAGAAAGAAAACGACTTATATTTCTTCGTGAGTGC CGCCTG	3005	
Qy	726	GCACCTCTGAGGAGTATATAATTAACACCACTTTTACAGCTAGACATCTTTTGTAGAA	785	
Db	3006	GCACCTCTGAGGAGTATATAATTAACACCACTTTTACAGCTAGACATCTTTTGTAGAA	3065	
Qy	786	AAGGCAATCGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCAATTAAGAGACAACCTCAC	845	
Db	3066	AAGGCAATCGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCAATTAAGAGACAACCTCAC	3125	
Qy	846	AATTAATGTAATAAGTGTGATTTATGCCCTTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC	905	
Db	3126	AATTAATGTAATAAGTGTGATTTATGCCCTTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC	3185	
Qy	906	CCAGCATCCCGACTCTCTCCCMYTAATAAGAACCCCCCTTCAACCCAAATGGTCCAA	965	
Db	3186	CCAGCATCCCGACTCTCTCCCAACTAAATAAGAGCCCCCTTCAACCCAAATGGTCCAA	3245	
Qy	966	AAGGAGATGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATCCCAATTTATGA	1025	
Db	3246	AAGGAGATGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATCCCAATTTATGA	3305	
Qy	1026	CCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCAGCCAGAGTGCATGTGCTTTTTC	1085	
Db	3306	CCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCAGCCAGAGTGCATGTGCTTTTTC	3364	
Qy	1086	TCTCCAGACTTAAGCAATAAACAAGACTTAGGTAAATTCCTCAGATAAATCTGATGG	1145	
Db	3365	TCTCCAGACTTAAGCAATAAACAAGACTTAGGTAAATTCCTCAGATAAATCTGATGG	3424	
Qy	1146	CTATATTGRTTTTAAACAAGGGTTAGGCAAAATCTTTTGATCTCAGATGGAGATATATA	1205	
Db	3425	CTATATTGRTTTTAAACAAGGGTTAGGCAAAATCTTTTGATCTCAGATGGAGATATATA	3483	
Qy	1206	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCACCAATCTGCAGCCT	1265	
Db	3484	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCACCAATCTGCAGCCT	3543	
Qy	1266	GAGGTTTGGGATCTCTGATCTCAGTCAAGTCAATGGATGANGATGACAACAAGG	1325	
Db	3544	GAGGTTTGGGATCTCTGATCTCAGTCAAGTCAATGGATGANGATGACAACAAGG	3600	
Qy	1326	AAAGANAATGATTTCCCAAGCCAGCAGCAGTTCCTCAGTCTASACCTTCATTGGGGAC	1385	
Db	3601	AAAGANAATGATTTCCCAAGCCAGCAGCAGTTCCTCAGTCTASACCTTCATTGGGGAC	3657	
Qy	1386	ACAGAAATCAGTAACATGGGAGATTTGGTCTGCAGACATTTTGCTAACTTTGTGTCTASAA	1445	



QY 2525 GTACCCCTCAGACCCCAAGGCCAACAAGGATTCCAAAAGATTGTTAAAGGACTTAAAGGC 2584  
DB GTACCCCTCAGACCCCAAGGCCAACAAGGACTTCCAAAAGATTGTTAAAGGACTTAAAGGC 4847  
QY 2585 CCAAGGCTTAGTAAACCATGATCAATCCCTCGCAGTAAATCCGTAGTGGATTGAGGAGG 2644  
DB CCAAGGCTTAGTAAACCATGATCAATCCCTCGCAGTAAATCCGTAGTGGATTGAGGAGG 4900  
QY 2645 CACAGAAACCCAGTGGACAGTGGAGGGTTAGTGGAGAGTCTCAGAGATTATCAATGGAGGC 2704  
DB TACAGAAACCCACAGACAGTGGGA--GGTTAGTGGCAAGATCTCAGGATTATCAAT--GAGGC 4958  
QY 2705 GGTGTCTCTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATACCAGA 2764  
DB TGTGTCTCTTATAGCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATACCAGA 5018  
QY 2765 GGAACAGAGTGGTTTACASTCCTCGACCTTMAGATGCTTCTTCTGTCATCCCTGTACA 2824  
DB GGAACAGAGTGGTTTACASTCCTCGACCTTMAGATGCTTCTTCTGTCATCCCTGTACA 5078  
QY 2825 TCCTGACTCTCAATTCCTGTTTGGCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 2884  
DB TCCTGACTCTCAATTCCTGTTTGGCTTTGAAGATACTTCAAACCCARCATCTCAACTCAC 5138  
QY 2885 CTGAGACTTTTACCCCAAGGTTTCAAGGATAGTCCCATCTATTTGGCCAGGCATTAGC 2944  
DB CTGAGACTTTTACCCCAAGGTTTCAAGGATAGTCCCATCTATTTGGCCAGGCATTAGC 5198  
QY 2945 CCAAGACTTGAGYCAATYMTACATCTGGACACTCTGTGCTTCRTGTAGTGGATGATT 3004  
DB CCAAGACTTGAGYCAATYMTACATCTGGACACTCTGTGCTTCRTGTAGTGGATGATT 5256  
QY 3005 ACTTTTGGYGCYRTTCAGAAACCTTGTCATCAAGCCACCCCAAGCRCTCTTMAATTT 3064  
DB ACTTTTGGYGCYRTTCAGAAACCTTGTCATCAAGCCACCCCAAGCRCTCTTMAATTT 5316  
QY 3065 CCTCCYACCTGTGGCTCAGGTTTCCAAACSBARGCTCARTCTCTCTCACAGCAGT 3124  
DB CCTCCYACCTGTGGCTCAGGTTTCCAAACSBARGCTCARTCTCTCTCACAGCAGT 5376  
QY 3125 TAAATACCTTAGRCCTAAATATCAAGAGCCACAGGCCCTTCAGTGAAGGAAATCA 3184  
DB TAAATACCTTAGRCCTAAATATCAAGAGCCACAGGCCCTTCAGTGAAGGAAATCA 5377  
QY 3185 GCCTATCTGGCTTATCTCATFYCAAAACCCCTAAAGCAACTAAGRRRTTCTTGGCT 3244  
DB GCCTATCTGGCTTATCTCATFYCAAAACCCCTAAAGCAACTAAGRRRTTCTTGGCT 5492  
QY 3245 AAYAGGYTCTCCGAAATGATTTCCCGAGTWTGGCBAAATAGCCAGGYCATTAWATA 3304  
DB AAYAGGYTCTCCGAAATGATTTCCCGAGTWTGGCBAAATAGCCAGGYCATTAWATA 5493  
QY 3305 CASTAATTAAGGAACTCAGAAAGCAATPACCCATTTARTAGATGGAYAMCTGAAGYMR 3364  
DB CASTAATTAAGGAACTCAGAAAGCAATPACCCATTTARTAGATGGAYAMCTGAAGYMR 5551  
QY 3365 AAGTGGCTTTCCAGGCCCCCTAAGAAAGGCTTTAAACCCCAAGYCCAGTGTGTTAAGYTTGCC 3424  
DB AAGTGGCTTTCCAGGCCCCCTAAGAAAGGCTTTAAACCCCAAGYCCAGTGTGTTAAGYTTGCC 5611  
QY 3425 AACRGGGCAAGACTTTTSTYATATVTCACAGAAAAAACAGAAAYAGTCTRRGAGTCC 3484  
DB AACRGGGCAAGACTTTTSTYATATVTCACAGAAAAAACAGAAAYAGTCTRRGAGTCC 5611  
QY 3485 TTACACAGTCCRAGGGYAGCTTTGCAACCCYRTGCRVACCTGASTAAGGAAAYTGATG 3544  
DB TTACACAGTCCRAGGGYAGCTTTGCAACCCYRTGCRVACCTGASTAAGGAAAYTGATG 5717  
QY 3545 TAGTGGCAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTACAGTYKTAGTAT 3604  
DB TAGTGGCAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTACAGTYKTAGTAT 5777  
QY 3604 TAGTGGCAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTACAGTYKTAGTAT 5836

QY 3605 CTGAAGCAGTTAAATTAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
DB CTGAAGCAGTTAAATTAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 5896  
QY 3665 AYRGCATACTCACTGTCTAAAGGAGACTTGTGGCTGTGACAGCAACCTGTTTAAATATTC 3724  
DB AYRGCATACTCACTGTCTAAAGGAGACTTGTGGCTGTGACAGCAACCTGTTTAAATATTC 5956  
QY 3725 AGGCTCTATTACTTGAARGGCCAGTGTGACACTGTGACACTTGTGCAACTCTTAACCCAG 3784  
DB AGGCTCTATTACTTGAARGGCCAGTGTGACACTGTGACACTTGTGCAACTCTTAACCCAG 6016  
QY 3785 YCNCATTTCTTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAATTTCTC 3844  
DB YCNCATTTCTTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAATTTCTC 6076  
QY 3845 AAACCTATGCACTCGAGGGGACCTTCTTAGAGTTCCTTGTGACTGATCCYGCCTTCAAC 3904  
DB AAACCTATGCACTCGAGGGGACCTTCTTAGAGTTCCTTGTGACTGATCCYGCCTTCAAC 6135  
QY 3905 TTGTATCTGATGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGACGTG 3964  
DB TTGTATCTGATGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGACGTG 6195  
QY 3965 GTGAGTGAATGAATTTGAAAGTAACTCCCTCACTCCAGGAACTAGTGTGCTVAGCTR 4024  
DB GTGAGTGAATGAATTTGAAAGTAACTCCCTCACTCCAGGAACTAGTGTGCTVAGCTR 6255  
QY 4025 GCAGAACTAATAGCCCTCACCTGGGCACTAGAAATTTAGGAGAAAGGAAAGGCAATATA 4084  
DB GCAGAACTAATAGCCCTCACCTGGGCACTAGAAATTTAGGAGAAAGGAAAGGCAATATA 6315  
QY 4085 TATCAGACTCTARTATGCTTACTAGTCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4144  
DB TATCAGACTCTARTATGCTTACTAGTCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 6375  
QY 4145 AGGAAATCTTAACTTTCYAGGAAACCTTATCAGAACTAGGAGAAAGGCAATATA 4204  
DB AGGAAATCTTAACTTTCYAGGAAACCTTATCAGAACTAGGAGAAAGGCAATATA 6435  
QY 4205 TTAATGCGWGTACAGAAACCTTARAGAGTGGAGTCTTACACTGCTGCGGGTCAATCANAA 4264  
DB TTAATGCGWGTACAGAAACCTTARAGAGTGGAGTCTTACACTGCTGCGGGTCAATCANAA 6495  
QY 4265 GGAAGAAAGGAAATTAAGAAAGTGTGCAAGCAATATTTGAAGCAAAAGAGCTGCA 4324  
DB GGAAGAAAGGAAATTAAGAAAGTGTGCAAGCAATATTTGAAGCAAAAGAGCTGCA 6555  
QY 4325 AGGAGGAGCCCTCCATTTAGAAATGCTTATTAACCTTCCCTTAGTATAGGTAATCCCTTC 4384  
DB AGGAGGAGCCCTCCATTTAGAAATGCTTATTAACCTTCCCTTAGTATAGGTAATCCCTTC 6615  
QY 4385 CGGAAACCAAGCCCTCAGTCTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG--CAG 4443  
DB CGGAAACCAAGCCCTCAGTCTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG--CAG 6675  
QY 4444 TTTTCTCCCTCGGAGCGGTTAGCCACTGGAAGAGGAAATACTTTTTCCTGCTGCAACTAT 4503  
DB TTTTCTCCCTCGGAGCGGTTAGCCACTGGAAGAGGAAATACTTTTTCCTGCTGCAACTAT 6735  
QY 4504 CCAATGGAATTAATTAACCCCTTCAACCTTCACTTACCTTACCTTACCTTACCTTACCTT 4563  
DB CCAATGGAATTAATTAACCCCTTCAACCTTCACTTACCTTACCTTACCTTACCTTACCTT 6795  
QY 4564 TCARATGCGCAAAATCATTTTACTTGGACAGGCTTTTCAAAAATCTAACAAGCATATKT 4623  
DB TCARATGCGCAAAATCATTTTACTTGGACAGGCTTTTCAAAAATCTAACAAGCATATKT 6855  
QY 4624 CAGGCTCTGTGAATGTGCCARARAAATAATCCCTGCTTATCCGCAAGCTCTTTCAGG 4683  
DB CAGGCTCTGTGAATGTGCCARARAAATAATCCCTGCTTATCCGCAAGCTCTTTCAGG 6915  
QY 4684 ARAACAAARAAACAGGCCATTACCTGTRARARAACTAGGCAATCTGATTTTACCCCAAGCCC 4743



Db	6916	AGAAACAAAGAACAGGCCATTACCTCGAGAGAGCTGGCAACTGATTTTACCCACAGCCCC	6975	TTCACTGCCACACCCCATATGCCCCGAACTGTATATCACTCTGCCACTCTTTGCATGCAT	8054
QY	4744	AAACCTCAGGGATTTCACTAGTCTACTAGTCTGGGTATATCTTTTACGGGTGGGCARAG	4803	GCAATACTCATTTATTTGGACAGGAAAAATGATTAATCTCTAGTTGTCTCGAGACTTTGGA	5883
Db	6976	AAACCTCAGGGATTTCACTAGTCTACTAGTCTGGGTATATCTTTTACGGGTGGGCAGAG	7035	GCAATACTCATTTATTTGGACAGGAAAAATGATTAATCTCTAGTTGTCTCGAGACTTTGGA	8114
QY	4804	GCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTATTAAGGACACTAGTTTCATGAATA	4863	GTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGGTGGAGTTCAA	5943
Db	7036	GCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTATTAAGGACACTAGTTTCATGAATA	7095	GTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGGTGGAGTTCAA	8174
QY	4864	ATTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGCATATGAGCCCTGCTTTCCAGGCC	4923	GATCAGGCAAGCAAAAAATGTAAGAGTAAATCTCCCAACTCATCCSGGTACATGGC	6003
Db	7096	ATTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGCATATGAGCCCTGCTTTCCAGGCC	7155	GATCAGGCAAGCAAAAAATGTAAGAGTAAATCTCCCAACTCATCCSGGTACATGGC	8234
QY	4924	ACAGTAACCCAGGGAGTATCCAGGGCTTAGGTATACGATATCACTTACACTGCGCCTGA	4983	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7156	ACAGTAACCCAGGGAGTATCCAGGGCTTAGGTATACGATATCACTTACACTGCGCCTGA	7215	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	4984	AGGCCACAGTCTTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAAAA	5043	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7216	AGGCCACAGTCTTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAAAA	7275	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5044	AAGCAAAACCCAGAAACCCACCTCACATGCGCTGTCTGTGCTATAGCTTTAAAAAGA	5103	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7276	AAGCAAAACCCAGAAACCCACCTCACATGCGCTGTCTGTGCTATAGCTTTAAAAAGA	7335	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5104	ATCTGCAACTTTCCCAAAAGCAGGACTTAGGCCATACGAAATGCTGTATGGAAGGCC	5163	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7336	ATCTGCAACTTTCCCAAAAGCAGGACTTAGGCCATACGAAATGCTGTATGGAAGGCC	7395	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5164	TTCAATACCAATGACCTTGTGCTTGAACCAAGCAGCAACTTAGTTGACAGATCAACCT	5223	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7396	TTCAATACCAATGACCTTGTGCTTGAACCAAGCAGCAACTTAGTTGACAGATCAACCT	7455	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5224	CTTTAGCCAAATATCAACAGTTCTTAAACATTTACAGGAACCTATCCCTGAGAGAGG	5283	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7456	CTTTAGCCAAATATCAACAGTTCTTAAACATTTACAGGAACCTATCCCTGAGAGAGG	7515	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5284	GAAAGAACTATTCCACCCWGTGACATGTTATAGTCAAGTCCCTTCYCTCAATTTCCC	5343	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7516	GAAAGAACTATTCCACCCWGTGACATGTTATAGTCAAGTCCCTTCYCTCAATTTCCC	7575	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5344	CATCCCTAGATACATCTCGGGAAGGACCCCTACCCAGTCACTTTATYACCCCAACTGCGG	5403	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7576	CATCCCTAGATACATCTCGGGAAGGACCCCTACCCAGTCACTTTATYACCCCAACTGCGG	7635	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5404	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCCTGATGCTGC	5463	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7636	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCCTGATGCTGC	7695	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5464	CMAAGGAACTGAAATTCAGGAGACAAACCTAGCTATTCTCTGTAACCTCTAGAGGAT	5523	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7696	CMAAGGAACTGAAATTCAGGAGACAAACCTAGCTATTCTCTGTAACCTCTAGAGGAT	7755	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5524	TGCGCTGCTCTTCAAAACAAACAGGAGAAAGTAACTAAATCATAAATCCCCCATG	5583	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7756	TGCGCTGCTCTTCAAAACAAACAGGAGAAAGTAACTAAATCATAAATCCCCCATG	7814	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5584	GSCTCGCTTATCATATTTTCTCTKTATGTTSTTTTACCTTTTCACTCTCACTGCA	5643	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7815	GSCTCGCTTATCATATTTTCTCTKTATGTTSTTTTACCTTTTCACTCTCACTGCA	7874	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5644	CCCCCTCCATGCGCTGATGACAGTACCTCCCTYACCMAGAGTTTCTATGGAGATG	5703	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7875	CCCCCTCCATGCGCTGATGACAGTACCTCCCTYACCMAGAGTTTCTATGGAGATG	7934	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5704	CAGCGTCCCGAAATATTTGATGCCCCATTCGTATAGGAGTCTTTSAAAGGAAACCCCAAC	5763	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
Db	7935	CAGCGTCCCGAAATATTTGATGCCCCATTCGTATAGGAGTCTTTSAAAGGAAACCCCAAC	7994	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294
QY	5764	TTCACTGCCACACCCATATGCCCCGAACTGCTATCACTCTGCCACTCTTTGTCGAT	5823	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCT	8294

QY	6904	AGCCATGATGATCCCTGGATTCCTCCCTTCTTAGGACCTCTAGCAGCTATATATTGCTA	6963
Db	9135	AGCCAAATGATGATCCCTGGATTCCTCCCTTCTTAGGACCTCTAGCAGCTATATATTGCTA	9194
QY	6964	CTCCTCTTTGGACCTCTATCTTTTACCTCTCTTTGTTAACTTTGCTCTCTCCAGAAATCGAA	7023
Db	9195	CTCCTCTTTGGACCTCTATCTTTTAACTCTCTTTGTTAACTTTGCTCTCTCCAGAAATCGAA	9254
QY	7024	GCTGTAAACTACAAATGAGGCCAGATGCGATGCACTTAAGATCAATACCGCAGACCC	7083
Db	9255	GCTGTAAACTACAAATGAGGCCAGATGCGATGCACTTAAGATCAATACCGCAGACCC	9314
QY	7084	CTGGACCGCTGATGAGCCAGATCTGATGTTATGATCATCAAAAGGCACCCCTCTGAG	7143
Db	9315	CTGGACCGCTGATGAGCCAGATCTGATGTTATGATCATCAAAAGGCACCCCTCTGAG	9374
QY	7144	GAATCTCAGCTGACCAACTCTACTAGCCCAATTCAGCAGGAGCAGTTAGAGCGGT	7203
Db	9375	GAATCTCAGCTGACCAACTCTACTAGCCCAATTCAGCAGGAGCAGTTAGAGCGGT	9434
QY	7204	STCGGCCAACTCCGCCAAGACACTTAGTTTCTGTTGAGATGGGGACTGAGAGAC	7263
Db	9435	STCGGCCAACTCCGCCAAGACACTTAGTTTCTGTTGAGATGGGGACTGAGAGAC	9494
QY	7264	AGGACTAGCTGATTTCTCCTAGCTGATTAAGATCCYTAAGCTTAGTGGGAGGTGACC	7323
Db	9495	AGGACTAGCTGATTTCTCCTAGCTGATTAAGATCCYTAAGCTTAGTGGGAGGTGACC	9554
QY	7324	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGVTCAACCTGACCAATCAGAGAGCTC	7383
Db	9555	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGVTCAACCTGACCAATCAGAGAGCTC	9614
QY	7384	ACTAAATGCTAATTAGGCAAGACAGGAGGTAAGAATACCCCAATCATYTAATTGCMTG	7443
Db	9615	ACTAAATGCTAATTAGGCAAGACAGGAGGTAAGAATACCCCAATCATCTATTGSCCTG	9674
QY	7444	AGACACAGCAGAGGAGCAATGATCGGGATATAACCAAGTCTGAGCGCGCAACGG	7503
Db	9675	AGACACAGCAGAGGAGCAATGATCGGGATATAACCAAGTCTGAGCGCGCAACGG	9734
QY	7504	CAACCCCTTTGGTGGTCCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCACCTCT	7563
Db	9735	CAACCCCTTTGGTGGTCCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCACCTCT	9794
QY	7564	ATTAATCTTGCACTGCTC 7582	
Db	9795	ATTAATCTTGCACTGCA 9813	
RESULT 6			
AY101585			
LOCUS	AY101585	10222 bp	DNA linear PRI 11-FEB-2004
DEFINITION	Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVW1 locus, allele B, complete sequence.		
ACCESSION	AY101585		
VERSION	AY101585.1	GI:37544405	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 10222)		
AUTHORS	Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.		
TITLE	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
PUBMED	14757826		
REFERENCE	2 (bases 1 to 10222)		
AUTHORS	Mallet, F., Bouton, O. and Oriol, G.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France		
FEATURES	Location/Qualifiers		
source	1..10222		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/isolate="79"		
	/isolation_source="PBM of individual 79"		
	/db_xref="taxon:9606"		
	/chromosome="7"		
	/map="7q21-q22"		
	/sex="male"		
	/cell_type="PBM"		
	/note="caucasian"		
	isolated by PCR; ERVW1 locus, allele B		
	endogenous_virus: HERV-W"		
LTR	1..780		
mRNA	join(248..811,2887..9813)		
	/note="putative mRNA transcript 1; gag-pol pseudogene"		
mRNA	join(248..811,7585..9813)		
	/product="envelope glycoprotein"		
	/note="putative mRNA transcript 2"		
mRNA	join(248..811,9248..9813)		
	/note="putative mRNA transcript 3; no coding region"		
CDS	7812..9428		
	/note="syncytin"		
	/codon_start=1		
	/product="envelope glycoprotein"		
	/protein_id="AA068164.1"		
	/db_xref="GI:37544406"		
	/translation="MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYQEFWRMQR		
	GNIDAPSRSLSGKPTFTTAHTHPRNCYHSATLCHMANTHYTGKMLNPSGGLGV		
	TVCTYFTQMGSDGGVQDOAREKHVEISQLTVRHGTSPYKGLDLXKLHETLR		
	HTRVLSFNTLTGLHEVSAQNPNCWICLPNRPYVSIIPVEOMNFSFEINTSV		
	LVGPLVNLNLTHTSVNLTQKFSNTVTTNSQCIKRWTPPTQIVCLPSGIFPVCTSA		
	YRLNGSESNCFSLVLPVPTIYEQDLISYVTSKPNKRVPILPFVIGAGVLGALG		
	TGCIGITITQFYKLSQELNGDMRVADSLVTDQNLNSLAADVVLQNRADLLTAE		
	RGGCLFLGEBCCYVNGSIVTEKVEIRDIQRABEELNTPGWLLSOMMPWILP		
	FLGPLAAIILLFLGPCIFNLLVNFVSRIEAVKLQWEPKMQSKTKIYRRFLDRPAS		
	RSDVNDIKGTPPEISAAQPLLRNSAGSS"		
	9487..10222		
LTR	10222		
ORIGIN			
Query Match	86.1%; Score 6448; DB 5; Length 10222;		
Best Local Similarity	95.7%; Pred. No. 0;		
Matches	6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;		
QY	606	GATGGAAACCTTCCCGCAAGACAAAACGCCCTTAAGACGTATTCTGGAATAATTGGGA	665
Db	2886	GATGGAAACCTTCCCGCAAGACAAAACGCCCTTAAGACGTATTCTGGAATAATTGGGA	2945
QY	666	MCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTCTGAGTCCGCTG	725
Db	2946	CCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTCTGAGTCCGCTG	3005
QY	726	GCATCTCTGAGGAGTATAAATTATAACACCATCTTACAGTACAGTCTTTTCTAGAA	785
Db	3006	GCATCTCTGAGGAGTATAAATTATAACACCATCTTACAGTACAGTCTTTTCTAGAA	3065
QY	786	AAGCAATGAGTGAAGTGCATAGTACAACTTTCTTTTTCATTAGAGCAACTCAC	845
Db	3066	AAGCAATGAGTGAAGTGCATAGTACAACTTTCTTTTTCATTAGAGCAACTCAC	3125
QY	846	AATTATGTAAGTGTGATTATGCTCCCTCAGAGAGCTTCTCAGAGTCTACCTCCCTATC	905
Db	3126	AATTATGTAAGTGTGATTATGCTCCCTCAGAGAGCTTCTCAGAGTCTACCTCCCTATC	3185
QY	906	CCAGCATCCCGACTCTCTTCCCCAMYTAAATAGGACCCCTTCAACCAATAGTCCAA	965
Db	3186	CCAGCATCCCGACTCTCTTCCCCAMYTAAATAGGACCCCTTCAACCAATAGTCCAA	3245
QY	966	AAGGAGATAGACAAAAGGTAACAGTGAACCAAGAGTGCATATATATCCCAATTATGA	1025

Db	3246	AAGGAGATAGACAAAGGGTAAACAGATGAACAAAGAGTGCCAAATATTTCCCAATATGA	3305	
Qy	1026	CCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCCAAGCAGAGTGCAATGTGCVTTTTYY	1085	
Db	3306	CCCT-CCAAGCAGTGGGAGGAAGAAATTCGGCCCAAGCAGAGTGCAATGTGCVTTTTTC	3364	
Qy	1086	TCTCCAGACTAAAGCAATAAATAACAGACTTATAGTAAATTTCTCAGATAAYCTGTATGG	1145	
Db	3365	TCTCCAGACTAAAGCAATAAATAACAGACTTATAGTAAATTTCTCAGATAAACCTGTATGG	3424	
Qy	1146	CTATATTGRTGTTTTACAAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGATATA	1205	
Db	3425	CTATATTGATGTTTTACAAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGATATA-A	3483	
Qy	1206	TGTCACCTGCTAAATCAGACATAACCCCAATGAGAGAAGTGCCACCAATAACTGCAGCCT	1265	
Db	3484	TGTCACCTGCTAAATCAGACATAACCCCAATGAGAGAAGTGCCACCAATAACTGCAGCCT	3543	
Qy	1266	GAGGTTTTGGCGATCTCTGTGATCTCAGTCAGGTCATATGATNGGATGACAAAGAGG	1325	
Db	3544	GAGGTTTTGGCGATCTCTGTGATCTCAGTCAGGTCATATGATNGGATGACAAAGAGG	3600	
Qy	1326	AAAGANAATGATCCCAAGCCAGCAGCAGTGCTCCAGTCTCTASACCCCTATTGGGGAC	1385	
Db	3601	AAAGANAATGATCCCAAGCCAGCAGCAGTGCTCCAGTCTCTASACCCCTATTGGGGAC	3657	
Qy	1386	ACAGAAATCAGTAACATGGGAGATTTGGTGTGACAGACATTTGCTTAATCTGTGTCTASAA	1445	
Db	3658	ACAGAAATCAGTAACATGGGAGATTTGGTGTGACAGACATTTGCTTAATCTGTGTCTASAA	3716	
Qy	1446	GGACTAAGGAAATCAGGAGAAATCTATGAAATTTACTCAATGATGTCACCAATACACA	1505	
Db	3717	GGACTAAGGAAATCAGGAGAAATCTATGAAATTTACTCAATGATGTCACCAATACACA	3775	
Qy	1506	GGGGAAGGAGAAATCTACTCCCTTTCTGGAGAGACTAAGGGAGCAATTGAGGAAGC	1565	
Db	3776	GGGGAAGGAGAAATCTACTCCCTTTCTGGAGAGACTAAGGGAGCAATTGAGGAAGC	3834	
Qy	1566	GTGCTCTCTGTACCTGACTCTCTGAAGGCCAACTTAATCTTAAAGCGTAAGTTTATCA	1625	
Db	3835	GTGCTCTCTGTACCTGACTCTCTGAAGGCCAACTTAATCTTAAAGCGTAAGTTTATCA	3894	
Qy	1626	CTCAGTCAGTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCGTAGGCCCGGAGCAAA	1684	
Db	3895	CTCAGTCAGTGCAGACATTAGAAAAAACTTCAAAAGTCTGCGTAGGCCCGGAGCAAA	3954	
Qy	1685	ACTTAGAAACCTATTGAACTTGCCACATCTGCTTTTATATATAGAGATCAGAGGAGC	1744	
Db	3955	ACTTAGAAACCTATTGAACTTGCCACATCTGCTTTTATATATAGAGATCAGAGGAGC	4014	
Qy	1745	AGCGGAAACAGGACAAACGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGG	1804	
Db	4015	AGCGGAAACAGGACAAACGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGG	4074	
Qy	1805	CAAGTGACCTTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAATTTGAATGCCATATAGG	1864	
Db	4075	CAAGTGACCTTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAATTTGAATGCCATATAGG	4134	
Qy	1865	GCTTGCTTCAGTGGGCTCTACAGGACATTTAAAAAGATTTGTCAGTAGAGTAAG	1924	
Db	4135	GCTTGCTTCAGTGGGCTCTACAGGACATTTAAAAAGATTTGTCAGTAGAGTAAG	4194	
Qy	1925	CCGCCCTCTGCTCATGCCCTTATTTCAAGGGAATCAGTGGAGGCCCACTGCCCCAGG	1984	
Db	4195	CCGCCCTCTGCTCATGCCCTTATTTCAAGGGAATCAGTGGAGGCCCACTGCCCCAGG	4254	
Qy	1985	GGACAAAGGCTTTTGTAGTCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGG	2044	
Db	4255	GGACAAAGGCTTTTGTAGTCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGG	4314	
Qy	2045	TGCTGGGCAAGGCCCATCCGATGCCATCACCTCTACAGAGCCCTGGGTATGCTTGACC	2104	

Qy	3185	GCCTATACTGGCTTTATCCTCATCYCAAAA	CCCTAAAGCAACTAAGRRRTTCTTGCCRT	3244
Db	5433	GCCTATACTGGCTTTATCCTCATCCAAAA	CCCTAAAGCAACTAAGGGATTCCTTGCCGT	5492
Qy	3245	AAYAGGYTTCGCGGAATWGGATTC	CCCGAGTWTGCGRAAATAGCCAGCYCATTAATA	3304
Db	5493	AATAGGTTTCTGCCGAATWGGATT	CCCAGGTATGCGAATAATAGCCAGGTCAATAATA	5551
Qy	3305	CASATAATTAAGGAACTCAGAAAGCC	CAATACCCATTTARTAAAGATGAYAMCTGAAGYMR	3364
Db	5552	CACATAATTAAGGAACTCAGAAAGCC	CAATACCCATTTAGTAGATGGCAACTGAAGTAG	5611
Qy	3365	AAGTGGCTTCCAGGCCCTTAAGAAG	GGCCTTAAACCAGYCCAGTGTTAAGYTTGCC	3424
Db	5612	AAGTGGCTTCCAGGCCCTTAAAGAAG	GGCCTTAAACCAGYCCAGTGTTAAGYTTGCC	5657
Qy	3425	AACGGGCAAGACTTTTSTYATYRT	CAACAGAAAAACAGRAAYAGCTCTGGAGTCC	3484
Db	5658	AACAGGGCAAGACTTTTCTCATATGT	CAAGCAAGAAATAGCTCTAGGAGTCC	5716
Qy	3485	TTACACAGRTCCRAGGAYAGCTTGC	CAACCCYRTGGCRYACCTGAATAAGAAAYTGATG	3544
Db	5717	TTACACAGATCCGAGGATGAGCTTGC	CAACCTGTGGCATACCTGACTAAGGAAATCGATG	5776
Qy	3545	TAGTGGCAAGGGTGTGRCYCATTTGT	TTATYGGGTAGTGGTGGCAGTAGCAGTGYKTAGTAT	3604
Db	5777	TAGTGGCAAGGGTGTGACCTCATTTG	TTATCGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT	5836
Qy	3605	CTGAAGCAGTTAAATAATAACAGGR	AGAGATCTTACTGTGGACATCTCATGAKGTGA	3664
Db	5837	CTGAAGCAGTTAAATAATAACAGGR	AGAGATCTTACTGTGGACATCTCATGATGTGA	5896
Qy	3665	AYRCATACTCACTCTAAAGGACACT	TGTGGCTGTGACAAACYGTTTACTTTAAATPTC	3724
Db	5897	ATGGCACTCACTGCTAAAGGACACT	TGTGGCTGTGACAAACCTGTTTACTTTAAATGTC	5956
Qy	3725	AGGCTCTATTACTTGAAAGGCCAGT	GTGCBACTGTGCACTTGTGCAACTCTTTAAACCAG	3784
Db	5957	AGGCTCTATTACTTGAAAGGGCAGT	GTGCBGACTGTGCACTTGTGCAACTCTTTAAACCAG	6016
Qy	3785	YCNCACTTCTTCACAGCAATGAAGA	AAAGATARAAYATACTGTCAACAARTAAATTTCTC	3844
Db	6017	CCACATTTCTTCACAGCAATGAAGA	AAAGATARAAYATACTGTCAACAGTAATTTCTC	6076
Qy	3845	AAACCTATGCCACTCGAGGGGACCT	TYTAGAGTTCCYTTGACTGTATCCYGAACCTTCAAC	3904
Db	6077	AAACCTATGCCACTCGAGGGGACCT	TYTAGAGTTCTCTTTGACTGTATCCCGAACCTCAAC	6135
Qy	3905	TTGTATCTGATGGAAGTTCCTTTGT	ATAGAAAGGACTTCGAAAGYGGGGTAGCAGTG	3964
Db	6136	TTGTATCTGATGGAAGTTCCTTTGT	ATAGAAAGGACTTCGAAAGYGGGGTAGCAGTG	6195
Qy	3965	GTCACTGATATGGAATYTTGAAAGT	TAATCCCTCACTCCAGGAACTAGTGTCTYAGCTR	4024
Db	6196	GTCACTGATATGGAATYTTGAAAGT	TAATCCCTCACTCCAGGAACTAGTGTCTCAGCTA	6255
Qy	4025	GCAGAACTAATAGCCYTCAYTKGGG	CACTAGAAATTAGGAGAGRAAAAAGGGYAAATATA	4084
Db	6256	GCAGAACTAATAGCCCTCACCTTGG	CACTAGAAATTAGGAGAGRAAAAAGGGCAAAATATA	6315
Qy	4085	TATACAGACTCTRATATGCTYACCT	AGTCTAGTCTTCCATGCCCATGMRGCAATATGSARAGAA	4144
Db	6316	TATACAGACTCTAATATATGCTTAC	TAGTCTCTCATGCCCATGAGCAATATGGAAGAA	6375
Qy	4145	AGGAAATTCCTAACTTCYVAGRG	GAACACCTATCAMAATCAGGAAGCCATTAGGARATTA	4204
Db	6376	AGGAAATTCCTTAATCTTCGAGAG	GAACACCTATCAAACTCAGGAAGCCATTAGGAATTA	6435
Qy	4205	TTATYGGCGWGTACGAAACCTTAR	AGAGTGTGAGTCTTACACTGCYGGGGTCAATCANAA	4264
Db	6436	TTATYGGCGTACGAAACCTTAAAG	AGTGTGCACTTACACTCGCGGGTCAATCAGAAA	6495

QY	4265	GGAAAGPAAAGGGAATAATASAAAGGAAYTGCCAGCAKATATTTGAAGGMAAAAGAGCTGCA	4322
DB	6496	GGAAAGGAAAGGGAATATAGAAAGAAAGTTCGCAAGCAGATATTTGAAGCGCAAAAGAGCTGCA	6555
QY	4325	AGGCAGACACCTCCATTTAGAAATGCTTTATTAACCTTCCCTTTAGTTATATAGGGTAAATCCCTTC	4384
DB	6556	AGGCAGACACCTCCATTTAGAAATGCTTTATTAACCAACCCCTAGTTATATAGGGTAAATCCCTTC	6615
QY	4385	CGGGAACCAAGCCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGCAGG-CAG	4443
DB	6616	CGGGAACCAAGCCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGCAGGACAG	6675
QY	4444	TTTTTCTCCCTCGGACGGTTAGGCACCTGAAGAAAGGGAATAATCTTTTGCTCCTGCAACTAT	4503
DB	6676	TTTTTCTCCCTCGGACGGCTAGCACA CTGAAGAAAGGGAATAATCTTTTGCTCCTGCAACTAT	6735
QY	4504	CCAAATGGAATTACTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCACCCA	4563
DB	6736	CCAAATGGAATTACTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCACCCA	6795
QY	4564	TCARATGGCCAAATCATTTATTTACTGGACAGGCGCTTTTCAAACCTATCAAGCARATAKT	4623
DB	6796	TCARATGGCCAAATCATTTATTTACTGGACAGGCGCTTTTCAAACCTATCAAGCAGATAGT	6855
QY	4624	CAGGCGCTGTGAAGTGTGCCARAPAAATAATCCCTGCGCTYATGCGCAAGCTCCTTCAGG	4683
DB	6856	CAGGCGCTGTGAAGTGTGCCAGAAATAATCCCTGCGCTYATGCGCAAGCTCCTTCAGG	6915
QY	4684	ARACAAARACAGGCCATTACCTGPARAARACTGCAACTCATTTATCCCAACAGCCC	4743
DB	6916	AGAACAAAGAACAGGCCATTACCTGGAAGAACTGGCAACTGATTTATCCCAACAGCCC	6975
QY	4744	AAACCTCAGGGATTTCACTACTACTAGTCTGGGTARATATCTTTACGGGTGGGCAAG	4803
DB	6976	AAACCTCAGGGATTTCACTACTACTAGTCTGGGTAGTACTTTACGGGTGGGCAAG	7035
QY	4804	GCCTTCCCCTGTAGGAAGAAAAGGCCAACAGAGTAAATAAGGCATCTAGTTTCATGAATA	4863
DB	7036	GCCTTCCCCTGTAGGAAGAAAAGGCCAACAGAGTAAATAAGGCATCTAGTTTCATGAATA	7095
QY	4864	ATTTCCAGATTCGAGACTTCCCGAGGCTTACGAGTGACAAATAGCCCTGCTTTCCAGGCC	4923
DB	7096	ATTTCCAGATTCGAGACTTCCCGAGGCTTACGAGTGACAAATAGCCCTGCTTTCCAGGCC	7155
QY	4924	ACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATAGCATATCACTTACACTGCGCCTGA	4983
DB	7156	ACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATAGCATATCACTTACACTGCGCCTGA	7215
QY	4984	AGGCCACAGTCTCAGGAAAGGTCGAAATAATGAATGAAAYACTCAAAGGACATCTTAAAA	5043
DB	7216	AGGCCACAGTCTCAGGAAAGGTCGAAATAATGAATGAAAYACTCAAAGGACATCTTAAAA	7275
QY	5044	AAGCAAAACCCAGGAAACCCACCTCAGTGGCTGYTCTGTGCTATAGCCTTAAAAAGA	5103
DB	7276	AAGCAAAACCCAGGAAACCCACCTCAGTGGCTGYTCTGTGCTATAGCCTTAAAAAGA	7335
QY	5104	ATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	5163
DB	7336	ATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	7395
QY	5164	TTCTATAACCAATGACTTTGTGCTTTGACCCAGACAGCACAATCTAGTTGCGACATCACT	5223
DB	7396	TTCTATAACCAATGACTTTGTGCTTTGACCCAGACAGCACAATCTAGTTGCGACATCACT	7455
QY	5224	CCTTAGGCAAAATATCAACAGTTCTTTAAAAACATTTACAAAGGAACCTATCTCCTGAGAAGG	5283
DB	7456	CCTTAGGCAAAATATCAACAGTTCTTTAAAAACATTTACAAAGGAACCTATCTCCTGAGAAGG	7515
QY	5284	GAAGAAGACTATTTCCACCCGWTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTTCCC	5343
DB	7516	GAAGAAGACTATTTCCACCCCTGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTTCCC	7575
QY	5344	CATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTTATATACCCCAACTCGG	5403

Db 7576 CATCCCTAGATACATCTCGGAGGACCTACCCAGTCATTTTATCTACCCCAACTGCGG 7635  
QY 5404 TTAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAAAATCCTGATCTGC 5463  
Db 7636 TTAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAAAATCCTGATCTGC 7695  
QY 5464 CAAAGGAACCTGAAATTCAGAGGAGACAAACGCTAGCTATTCTCTGAACTCTAGAGGATT 5523  
Db 7696 CAAAGGAACCTGAAATTCAGAGGAGACAAACGCTAGCTATTCTCTGAACTCTAGAGGATT 7755  
QY 5524 TGCGCTGCTCTTTCAACAAACACAGAGGAGAAAGTAACTAAATCATAAATCCCCCATG 5583  
Db 7756 TGCGCTGCTCTTTCAACAAACACAGAGGAGAAAGTAACTAAATCATAAAT-CCCCATG 7814  
QY 5584 GSCCTCCCTTATCAATATTTTCTCTKTATSTGTTSTTTTAYACCTSTTTCACTCTCATGCA 5643  
Db 7815 GCCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCTCTTTCACTCTCATGCA 7874  
QY 5644 CCCCTCCATGCCGCTGTATGACCACTAGCTAGCTCCCTTACOMAGAGTTTCTATGGAGATG 5703  
Db 7875 CCCCTCCATGCCGCTGTATGACCACTAGCTAGCTCCCTTACOMAGAGTTTCTATGGAGATG 7934  
QY 5704 CAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCTTSTAAGGGAACCCCAACC 5763  
Db 7935 CAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCTTSTAAGGGAACCCCAACC 7994  
QY 5764 TTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGTCAT 5823  
Db 7995 TTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGTCAT 8054  
QY 5824 GCBAATACTCATTTTGGACAGGAAATATGATTTAACTCTAGTGTCTGCTGGAGACTTGA 5883  
Db 8055 GCAAACTACTCATTTTGGACAGGAAATATGATTTAACTCTAGTGTCTGCTGGAGACTTGA 8114  
QY 5884 GTCACTGTCTGTGGACTTTTACCCCAAACTGGTATGTCGATGGGTGGAGTTCAA 5943  
Db 8115 GTCACTGTCTGTGGACTTTTACCCCAAACTGGTATGTCGATGGGTGGAGTTCAA 8174  
QY 5944 GATCAGGCAAGGAAAAATGTAAAGAGTAATCTCCAACTCAACCGGGGTACATGGC 6003  
Db 8175 GATCAGGCAAGGAAAAATGTAAAGAGTAATCTCCAACTCAACCGGGGTACATGGC 8234  
QY 6004 ACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGGTACCCAT 6063  
Db 8235 ACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGGTACCCAT 8294  
QY 6064 ACTCGCTGTAGCCCTATTTAATACCAACCTCACTGGGCTCCATGAGTCTCGGCCAA 6123  
Db 8295 ACTCGCTGTAGCCCTATTTAATACCAACCTCACTGGGCTCCATGAGTCTCGGCCAA 8354  
QY 6124 AACCCTACTAACTGTTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 6183  
Db 8355 AACCCTACTAACTGTTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 8414  
QY 6184 GTACCTGAACATGGAAACAACTTCAGCAAGAAATAAACCACTTCCTGTTTAGTAGGA 6243  
Db 8415 GTACCTGAACATGGAAACAACTTCAGCAAGAAATAAACCACTTCCTGTTTAGTAGGA 8474  
QY 6244 CCTCTGTTTCCAAATSTGGAATTAACCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
Db 8475 CCTCTGTTTCCAAATSTGGAATTAACCATACCTCAAACTCACCTGTGTAAATTTAGC 8534  
QY 6304 AATACTACATACAAACCAACTCCCAATGATCAGGTGGGTAACTCCTCCCAACAAATA 6363  
Db 8535 AATACTACATACAAACCAACTCCCAATGATCAGGTGGGTAACTCCTCCCAACAAATA 8594  
QY 6364 GTCTGCTACCTCAGGAAATATTTTGTGTGTGTACCTCAGCCTATCGTTGTTTGAAT 6423  
Db 8595 GTCTGCTACCTCAGGAAATATTTTGTGTGTGTGTACCTCAGCCTATCGTTGTTTGAAT 8654  
QY 6424 GGCCTTTAGATCTATGCTTCTCTCATTTCTAGTGGCCCTCATGCTACACT 6483

Db 8655 GGCCTTTAGATCTATGCTTCTCTCATTTCTTAGTGCCCTATGACCATCTTCACT 8714  
QY 6484 GAAACAAAGATTTATACAGTTATGTATATCTTAAGCCCGCAACAAAGAGTACCATCTTT 6543  
Db 8715 GAAACAAAGATTTATACAGTTATGTATCTTAAGCCCGCAACAAAGAGTACCATCTTT 8774  
QY 6544 CCTTTTGTATAGGAGCAGGAGTGTAGTGCATCTAGGTACTGGCAATGGCGGTATCACA 6603  
Db 8775 CCTTTTGTATAGGAGCAGGAGTGTAGTGCATCTAGGTACTGGCAATGGCGGTATCACA 8834  
QY 6604 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGTTC 6663  
Db 8835 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGTTC 8894  
QY 6664 GCGACTCCCTGCTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA 6723  
Db 8895 GCGACTCCCTGCTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA 8954  
QY 6724 AATCGAAGACTTTAGACTTGTCTAACCGCTGABAGAGGGGAAACCTGTTTATTTTAGGG 6783  
Db 8955 AATCGAAGACTTTAGACTTGTCTAACCGCTGABAGAGGGGAAACCTGTTTATTTTAGGG 9014  
QY 6784 GAAGAAATGCTGTATTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 9015 GAAGAAATGCTGTATTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 9074  
QY 6844 CSAGATCGAATCAACGTAAGCAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCTCTC 6903  
Db 9075 CGAGATCGAATCAACGTAAGCAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCTCTC 9134  
QY 6904 AGCCTATGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
Db 9135 AGCCTATGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 9194  
QY 6964 CTCCTCTTTGGACCTCTGATCTTTTACCTCTCTGTTTAACTTTGCTCTTCCAGAAATCGA 7023  
Db 9195 CTCCTCTTTGGACCTCTGATCTTTTAACTCTCTGTTTAACTTTGCTCTTCCAGAAATCGA 9254  
QY 7024 GCTGTRAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAAGATCTACCGCAGACCC 7083  
Db 9255 GCTGTRAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAAGATCTACCGCAGACCC 9314  
QY 7084 CTGGACCGGCTGTAGCCCAAGATCTGATGTTTAAATGATCAAAAGGCAACCCCTCTGAG 7143  
Db 9315 CTGGACCGGCTGTAGCCCAAGATCTGATGTTTAAATGATCAAAAGGCAACCCCTCTGAG 9374  
QY 7144 GAATCTCAGCTCAGCAACCTCTACTACGCCCAATTCAGCAGGAGCAGTTAGACGGT 7203  
Db 9375 GAATCTCAGCTCAGCAACCTCTACTACGCCCAATTCAGCAGGAGCAGTTAGACGGT 9434  
QY 7204 SGTCGGCCAAACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGATGGGGAGCTGAGAGAC 7263  
Db 9435 CGTCGGCCAAACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGATGGGGAGCTGAGAGAC 9494  
QY 7264 AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAAGGTGACC 7323  
Db 9495 AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAAGGTGACC 9554  
QY 7324 ACATCCACTTTTAAACACGGGCTTGCACCTTAGTGTACACCTGACCAATTCAGAGAGCTC 7383  
Db 9555 ACATCCACTTTTAAACACGGGCTTGCACCTTAGTGTACACCTGACCAATTCAGAGAGCTC 9614  
QY 7384 ACTAAATGCTAAATTTAGGCAAGACAGGAGGTAAGAAATACCAATCATYTTATTCMTG 7443  
Db 9615 ACTAAATGCTAAATTTAGGCAAGACAGGAGGTAAGAAATACCAATCATYTTATTCMTG 9674  
QY 7444 AGAGCAGCAGCAGGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCCGGAACCG 7503  
Db 9675 AGAGCAGCAGCAGGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCCGGAACCG 9734  
QY 7504 CAAACCCCTTTGGGTCCTCTCTGTTGATGAGGAGCTCTGTTTTCATGCTATTTCACTCT 7563  
Db 9735 CAAACCCCTTTGGGTCCTCTCTCTGTTGATGAGGAGCTCTGTTTTCATGCTATTTCACTCT 9794









Qy	4444	TTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGAAAAATATCTTTTGGCTGCAACTAT	4503
Db	34743	TTTTCTCCCTCGGGACGGCTAGCCACTGAAGAGGAAAAATCTTTTGGCTGCAACTAT	34802
Qy	4504	CCAAATGGAATTAATCTTAAACCCCTTCATCAAAACCTTTACATTAGGCATCGATAGCACCCA	4563
Db	34803	CCAAATGGAATTAATCTTAAACCCCTTCATCAAAACCTTTACATTAGGCATCGATAGCACCCA	34862
Qy	4564	TCARATGGCCAAATCATTTATTTACTGGACAGGSCCTTTTCAAAACTATCAAGACARATAKT	4623
Db	34863	TCAGATGGCCAAATCATTTATTTACTGGACAGGSCCTTTTCAAAACTATCAAGACAGATAGT	34922
Qy	4624	CAGGGCCTGTGAATGTGCCARARAAAATTAATCCCTGCGCTYATCCCAAGCTCTCTCAGG	4683
Db	34923	CAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCTGCGCTTATCGCAAGCTCTCTCAGG	34982
Qy	4684	ARAACAARAACAGGCCATTACCCCTGPARAARAACCTGGCAACTGATTTTACCCACNAGCCC	4743
Db	34983	AGAAACAAGAACAGGCCATTACCCCTGGAGAGAAGCTGGCAACTGATTTTACCCACNAGCCC	35042
Qy	4744	AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATACTTTTCAACGGGTGGGCARAG	4803
Db	35043	AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTAGATACTTTCAACGGGTGGGCAGAG	35102
Qy	4804	GCCTTCCCTGTAGGACAGAAAAAGGCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA	4863
Db	35103	GCCTTCCCTGTAGGACAGAAAAAGGCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA	35162
Qy	4864	ATTCCAGATTCGGACTTTCCCGAGGCTTACAGAGTGACATAGAGCCTGCTTTCAGGCC	4923
Db	35163	ATTCCAGATTCGGACTTTCCCGAGGCTTACAGAGTGACATAGAGCCTGCTTTCAGGCC	35222
Qy	4924	ACAGTAAACAGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	4983
Db	35223	ACAGTAAACAGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	35282
Qy	4984	AGGCCACAGTCTCCAGGAAAGTCCGAGAAATGAATGAAYAACTCAAAAGGCATCTAAAA	5043
Db	35283	AGGCCACAGTCTCCAGGAAAGTCCGAGAAATGAATGAAYAACTCAAAAGGCATCTAAAA	35342
Qy	5044	AAGCAAAACAGGAAACCCACCTCAATGGCTGYTCTGTTGGCTTATAGCCTTAAAAAGA	5103
Db	35343	AAGCAAAACCCAGGAAACCCACCTCAATGGCTGCTCTGTTGGCTTATAGCCTTAAAAAGA	35402
Qy	5104	ATCTGCAACTTTCCCCAAAAGCAGACGACTAGGCCCATCGAAATGCTGTATGGAAGGCC	5163
Db	35403	ATCTGCAACTTTCCCCAAAAGCAGACGACTAGGCCCATCGAAATGCTGTATGGAAGGCC	35462
Qy	5164	TTCTATAACCAATGACCTTGTGCTTGACCCCAAGACGCCAACTTAGTTTGACAGATCACCT	5223
Db	35463	TTCTATAACCAATGACCTTGTGCTTGACCCCAAGACGCCAACTTAGTTTGACAGATCACCT	35522
Qy	5224	CCTTAGCCAAATATCAACAAGTTCTTAAACATTACAAGGAACCTATACCTGTGAGAAGAGG	5283
Db	35523	CCTTAGCCAAATATCAACAAGTTCTTAAACATTACAAGGAACCTATACCTGTGAGAAGAGG	35582
Qy	5284	GAAGAAGACTTATCCACCCCHWGTGACATGGTATTAGTCAAGTCCCTTCCTCTATTTCCC	5343
Db	35583	GAAGAAGACTTATTCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCCTCTATTTCCC	35642
Qy	5344	CATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATATACCCCAACTGGG	5403
Db	35643	CATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATCTACCCCAACTGGG	35702
Qy	5404	TTAAAGTGGCTGGAGTCTGTGGATACATCACACTTTGAGTCAAAATCTCGGATCTGCG	5463
Db	35703	TTAAAGTGGCTGGAGTCTGTGGATACATCACACTTTGAGTCAAAATCTCGGATCTGCG	35762
Qy	5464	CAAGGAACCTGAAATTCAGGAGACAAACGCTAGCTATTTCCTGTGNAACCTCTAGAGGATT	5523
Db	35763	CAAGGAACCTGAAATTCAGGAGACAAACGCTAGCTATTTCCTGTGNAACCTCTAGAGGATT	35822
Qy	5524	TGCGCCTGCTCTTCAAAACAACACAGGAGGAAGTAACCTAAAAATCATAAATCCCCCATG	5583

Db	35823		TGGCGCTGCTCTTTCAAAACAACAACACGAGGAGGAAGTAACATAAAATCATAAAT-CCCCATG	35881
QY	5584	GSCCTCCCTTATCATATTTTTCTCTKTAAGTGTSTTTVACCCCTSTTTCACTCTCACTGCA	5643	
Db	35882	GCCTCCCTTATCATATTTTTCTCTTACTGTCTTTTACCCTCTTTCACTCTCACTGCA	35941	
QY	5644	CCCCCTCCATGCGCGCTGTATGACCACTAGTCTCCCTCYAACMAGAGTTTCTATGGAGAAATG	5703	
Db	35942	CCCCCTCCATGCGCGCTGTATGACCACTAGTCTCCCTTACCAAGAGTCTTCTATGGAGAAATG	36001	
QY	5704	CAGGCTCCCGGAAATATTCATGTGCCCATCGTATAGAGTCTTTSTAAGGAAACCCCAACC	5763	
Db	36002	CAGGCTCCCGGAAATATTCATGTGCCCATCGTATAGAGTCTTTCTTAAGGAAACCCCAACC	36061	
QY	5764	TTCACTGCCCACACCCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTTGCAATGCAT	5823	
Db	36062	TTCACTGCCCACACCCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTTGCAATGCAT	36121	
QY	5824	GCAATACTCATTTATTTGGACAGGAAAAATGATTAATCTCTAGTTGCTCTGGAGCACTTGA	5883	
Db	36122	GCAATACTCATTTATTTGGACAGGAAAAATGATTAATCTCTAGTTGCTCTGGAGCACTTGA	36181	
QY	5884	GTCACTGTCGTGTGGACTTACTTTCACCCAAACTGGTATGTCATGATGGGGTGGAGTTCAA	5943	
Db	36182	GTCACTGTCGTGTGGACTTACTTTCACCCAAACTGGTATGTCATGATGGGGTGGAGTTCAA	36241	
QY	5944	GATCAGCGAAGAGAAAACATGTAAAGAAAGTAATCTCCCAACTCACCGSGGTACATGGC	6003	
Db	36242	GATCAGCGAAGAGAAAACATGTAAAGAAAGTAATCTCCCAACTCACCGSGGTACATGGC	36301	
QY	6004	ACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	6063	
Db	36302	ACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	36361	
QY	6064	ACTGCGCTGTAAGCCTATTTTAATACCAACCTCACTGGGCTCCATGAGGTCYCGGCCCAA	6123	
Db	36362	ACTGCGCTGTAAGCCTATTTTAATACCAACCTCACTGGGCTCCATGAGGTCYCGGCCCAA	36421	
QY	6124	AACCTTACTAACTGTGTGGATATGCTCCCTCGGAACTTCARGCCATATGTTTCAATCCCT	6183	
Db	36422	AACCTTACTAACTGTGTGGATATGCTCCCTCGGAACTTCARGCCATATGTTTCAATCCCT	36481	
QY	6184	GTACCTGAAACAATGGGAACAACCTTCAGCACAGAAATAAACACACTTCOGTTTTAGTAGGA	6243	
Db	36482	GTACCTGAAACAATGGGAACAACCTTCAGCACAGAAATAAACACACTTCOGTTTTAGTAGGA	36541	
QY	6244	CCTCTGTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACCTGTGTGTAAATTTAGC	6303	
Db	36542	CCTCTGTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACCTGTGTGTAAATTTAGC	36601	
QY	6304	AATACTACATACACAACCAACTCCCAATGCATCAGTGGGTAACTCTCCCAACACAAATA	6363	
Db	36602	AATACTACATACACAACCAACTCCCAATGCATCAGTGGGTAACTCTCCCAACACAAATA	36661	
QY	6364	GTCTGCCCTACCTCAGGAATATTTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGAAT	6423	
Db	36662	GTCTGCCCTACCTCAGGAATATTTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGAAT	36721	
QY	6424	GGCTCTTTCAGAACTATGTGCTTCCTCTCATCTTATGTAGTCCCCCVATGRCATCTACACT	6483	
Db	36722	GGCTCTTTCAGAACTATGTGCTTCCTCTCATCTTATGTAGTCCCCCTATGACCACTTACACT	36781	
QY	6484	GAACAAGATTTATACAGTTATGTGCATATCTAAGCCCGCAACAAAGAGTACCCATCTTT	6543	
Db	36782	GAACAAGATTTATACAGTTATGTGCATATCTAAGCCCGCAACAAAGAGTACCCATCTTT	36841	
QY	6544	CTTTTCTGTATAGGACGAGGATGCTAGGTGCACTAGGTACTTGGCAATGGCGGTATCACA	6603	
Db	36842	CTTTTCTGTATAGGACGAGGATGCTAGGTGCACTAGGTACTTGGCAATGGCGGTATCACA	36901	
QY	6604	ACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGGTTC	6663	

Db	36902	ACCTCTACTCAGTTCTACTACAAATATCTCAAGAACTATAATGGGACATGGAACGGGTC	36961
Qy	6664	GCCGACTCCCTGGTCACTTGAAGATCAACTTAATCCCTAGCAGCAGTAGTCTTCRA	6723
Db	36962	GCCGACTCCCTGGTCACTTGAAGATCAACTTAATCCCTAGCAGCAGTAGTCTTCRA	37021
Qy	6724	AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAACCTGTTATTTTAGGG	6783
Db	37022	AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAACCTGTTATTTTAGGG	37081
Qy	6784	GAGAAATGCTGTTATTAATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAAT	6843
Db	37082	GAGAAATGCTGTTATTAATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAAT	37141
Qy	6844	CSAGATCGAATACAACTAGACGAGAGAGCTTGAACAACTGACCCCTGGGCGCTCCTC	6903
Db	37142	CGAGATCGAATACAACTAGACGAGAGAGCTTGAACAACTGACCCCTGGGCGCTCCTC	37201
Qy	6904	AGCRAATGATGCGCTTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTCCTA	6963
Db	37202	AGCRAATGATGCGCTTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTCCTA	37261
Qy	6964	CTCCTCTTTGGACCTGTATCTTTTACCTCTTGTAACTTTGTTCTCTCCAGAAATCGAA	7023
Db	37262	CTCCTCTTTGGACCTGTATCTTTTACCTCTTGTAACTTTGTTCTCTCCAGAAATCGAA	37321
Qy	7024	GCTGTAAATCAATGAG	7083
Db	37322	GCTGTAAATCAATGAG	37381
Qy	7084	CTGACCGGCTGTATGAGCCAGCATCTGATGTTAATGACATCAAGGCAACCTCCCTGAG	7143
Db	37382	CTGACCGGCTGTATGAGCCAGCATCTGATGTTAATGACATCAAGGCAACCTCCCTGAG	37441
Qy	7144	GAAATCTCAGCTGCACAACTCTACTACGCCCAATTTAGCAGAGAGAGAGAGAGAGAG	7203
Db	37442	GAAATCTCAGCTGCACAACTCTACTACGCCCAATTTAGCAGAGAGAGAGAGAGAGAG	37501
Qy	7204	SGTCGCGCAACCTCCCAACAGACACTAGATTTTCTGTTGAGATGGGAGCTGAGAGAC	7263
Db	37502	SGTCGCGCAACCTCCCAACAGACACTAGATTTTCTGTTGAGATGGGAGCTGAGAGAC	37561
Qy	7264	AGGACTAGCTGGATTTCTTAGGCTGATTAAGATCCCTTAAGCTAGCTAGGAGGTGACC	7323
Db	37562	AGGACTAGCTGGATTTCTTAGGCTGATTAAGATCCCTTAAGCTAGCTAGGAGGTGACC	37621
Qy	7324	ACATCCACCTTTAAACACCGGGCTTGCACCTTAACTAGTCAACCTGACCAATCAGAGAGCTC	7383
Db	37622	ACATCCACCTTTAAACACCGGGCTTGCACCTTAACTAGTCAACCTGACCAATCAGAGAGCTC	37681
Qy	7384	ACTAAATGCTAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	7443
Db	37682	ACTAAATGCTAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	37741
Qy	7444	AGAGCAGCAG	7503
Db	37742	AGAGCAGCAG	37801
Qy	7504	CAACCCCTTTGGGCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT	7563
Db	37802	CAACCCCTTTGGGCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT	37861
Qy	7564	ATTAATCTTGCACCTGCR	7582
Db	37862	ATTAATCTTGCACCTGCA	37880

RESULT 8  
HSAC000064  
LOCUS  
DEFINITION Human BAC clone RG083M05 from 7q21-7q22, complete sequence.  
AC000064  
AC000064.1 GI:1669369

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

HTG.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 56093)  
Pauley, A.  
The sequence of H. sapiens BAC clone RG083M05  
Unpublished (1996)  
2 (bases 1 to 56093)  
Waterston, R.  
Direct Submission  
Submitted (13-NOV-1996)  
Genome Sequencing Center  
Department of Genetics, Washington University  
St. Louis, MO 63108, USA  
e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:  
This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).  
VECTOR: pBelo  
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H\_RG083M05; actual end is at 56093 of H\_RG083M05

This clone contains STS sw351725.

FEATURES  
Source  
Location/Qualifiers  
1..56093  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="7"  
/map="7q21-7q22"  
/clone="H RG083M05"  
/clone\_lib="CITB-978SK-B"  
complement(838..1131)  
/rpt\_family="ALU"  
<1360..16971  
/gene="WUGSC:H RG083M05.1"  
join(<1360..1503,4181..4370,4587..4774,6422..6556,9483..9547,11631..11773,11864..12021,13131..13296,14885..14988,16349..16546,16837..16971)  
/gene="WUGSC:H RG083M05.1"  
/note="ATPase; strong similarity to paroxisome biosynthesis protein PAS1 (P1D:51172019); coded for by human cDNA C04279 (NID:61467530)"  
/codon\_start=1  
/protein\_id="PAB46346.1"  
/db\_xref="GI:1669371"  
/translation="KRLENIQKLEAFSEAVMMOPSVLLDLDLILAGLPAPSEHSPDARQCEILCNVINKKLDCKINKFTDLQHVAKETGGFVARDFTVLVRAHSLRQISITREKLVLITLDFQKALRGFLPASLRSLNLRKPDLDGWDKIGLHVRQILMDTIQPAKVLKKEKYPFLANLPIDRTGILLYPGPTGKTLGAGVIAHESRMFISVKGPELLSKYIGASEQAVRDIETRAQAAPKPCILFFDEFESIAPRRHDNTGVTDRVNVQ

```
LLTLDGVGLOGVYVLAATSRPDLIDPALLRPGRLDKVCYPPPDQVTVSYLSKTKO
QMLHSFLVSRLEINLVSDSLPLADDVLOHVASVTSFTGADLKALLYNAQLEALHG
KLLKMSILPDEKSPNNRYLYFGSSYESELNGTSDLSQCLSAFSSMTQDILPGVP
GKDQFSPQPTASQBGQELTQEQDQRADISIIKGYRSQSGEDSMNQPGPI
KTRLAISHLMTALGHTPRPSISEDDWKNFAEL"
/ complement (4948). .5130)
/rpt family="ALU"
/ complement (6581). .7133)
/rpt family="L1"
/ complement (7767). .8037)
/rpt family="ALU"
/ complement (8186). .8472)
/rpt family="ALU"
8473..8625
/ gene="WUGSC:H RG083M05.1"
/ note="match to human 3' EST H75782 (NID:g1049794), bases
287-444"
8841..9161
/ gene="WUGSC:H RG083M05.1"
/ note="match to human 5' EST H75921 (NID:g1050050), bases
21-348"
9481..9547
/ gene="WUGSC:H RG083M05.1"
/ note="match to human 5' EST N22627 (NID:g1130501), bases
276-343"
/ complement (12612). .12907)
/rpt family="ALU"
13670..13793
/ gene="WUGSC:H RG083M05.1"
/ note="match to human 5' EST H41382 (NID:g17434), bases
143-266"
13794..13877
/rpt family="ALU"
13878..13906
/ gene="WUGSC:H RG083M05.1"
/ note="match to human 5' EST H41382 (NID:g917434), bases
20-58"
13907..14104
/rpt family="ALU"
/ complement (14110). .14137)
/rpt family="L1"
/ complement (15618). .15907)
/rpt family="ALU"
17227..17522
/rpt family="ALU"
18667..19235
/ note="match to human fetal brain 5' EST D61494
(NID:g970409), bases 1-255, and to human 3' EST R07476
(NID:g759393)"
19550..19670
/rpt family="ALU"
21507..37303
/ note="similarity to various SS-RNA virus polyproteins;
pseudogene; region of matches and close matches to
multiple human ESTs, see R68740 (NID:g842257)"
37316..37489
/ inference="non-experimental evidence, no additional
details recorded"
/ note="Grail prediction, score = 80"
/ complement (38938). .39224)
/rpt family="ALU"
39225..39707
/ note="match to multiple human ESTs, see N30113
(NID:g1148633)"
39800..40085
/rpt family="ALU"
/ complement (40247). .40538)
/rpt family="ALU"
/ complement (40632). .40924)
/rpt family="ALU"
/ complement (42283). .42891)
/rpt family="ALU"
/ complement (45474). .45613)

/rpt family="ALU"
/ complement (45614). .45737)
/ note="match to human 3' EST H48898 (NID:g988738), bases
129-333"
/ note="match to multiple human ESTs, see N81064
(NID:g1243765), H48897 (NID:g988737), and M78831
(NID:g273146)"
/ complement (46107). .47026)
/ note="match to multiple human ESTs, see W37495
(NID:g1319089)"
/ complement (47027). .47318)
/rpt family="ALU"
/ note="match to multiple human ESTs, see W37495
(NID:g1319089)"
47898..48115
/ note="match to human 5' EST H62306 (NID:g1015138), bases
93-368"
/ complement (48116). .48405)
/rpt family="ALU"
/ complement (48406). .48584)
/ note="match to human 3' EST N29952 (NID:g1148472), bases
290-455, and 5' EST R12730 (NID:g765806)"
/ complement (48787). .49405)
/rpt family="ALU"
/ complement (49406). .49534)
/ note="match to human 3' EST R65794 (NID:g838432), bases
309-440"
/ complement (49638). .49672)
/rpt family="ALU"
/ complement (49674). .49890)
/ note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
/ complement (49698). .51806)
/ gene="WUGSC:H RG083M05.2"
/ complement (join (49698). .49888, 51575. .51806))
/ gene="WUGSC:H RG083M05.2"
/ note="coded for by human cDNAs W37389 (NID:g1319205),
R65891 (NID:g838529), R65794 (NID:g838432) and R65794
(NID:g838432)"
/ codon_start=1
/ protein_id="AAB46345.1"
/ db_xref="GI:16693370"
/ translation="MFVYFCGIIFCPPQGVYQIGDVSVVIDEQDGKPYAQIRGF
IQDYCEKSAALTWLPTLSRPDQFDPASVYIGPEDLPRKMEYLFVCHAPSVEFK
SRSPFPVTRPEKGIWTHVGTPTAIKESVANHL"
/ complement (51576). .51758)
/ gene="WUGSC:H RG083M05.2"
/ inference="non-experimental evidence, no additional
details recorded"
/ note="Grail prediction, score = 86"

Query Match 86.1%; Score 6448; DB 5; Length 56093;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGGAACCTGCCCGCAAGACAAAACGCCCTTAAGACGTATTCTGGAAATTTGGGA 665
Db 30953 GATGGGAACCTGCCCGCAAGACAAAACGCCCTTAAGACGTATTCTGGAAATTTGGGA 31012
QY 666 MCAATTTGACCTCAGACACTAAGAAAGAACGACTTATTTCTTCGAGTCCGCCCTG 725
Db 31013 CCAATTTGACCTCAGACACTAAGAAAGAACGACTTATTTCTTCGAGTCCGCCCTG 31072
QY 726 GCATCTCTGAGGAAGTATAAATTATAACACCATCTTACAGTACAGTCCTTTTAGTAA 785
Db 31073 GCATCTCTGAGGAAGTATAAATTATAACACCATCTTACAGTACAGTCCTTTTAGTAA 31132
QY 786 AAGCAATGGAGTGAAGTGCCATAAGTACAAAACCTTTCTTTTCATTAAAGACAACTCAC 845
Db 31133 AAGCAATGGAGTGAAGTGCCATAAGTACAAAACCTTTCTTTTCATTAAAGACAACTCAC 31192
QY 846 AATTATGTAAAAAGTGTATTATCGCTTACAGGAAGCCTTCAGAGTCTTACCTCCATATC 905
```



Qy	3065	CCTCGCVACCTGTGGCTACAGWGGTTTCCAAA	CSARARGCTCARCTCTGCTACAGCAGGT	3124
Db	33384	CCTCGCTACCTGTGGCTACATGGTTTCCAAA	CCAAAGGCTCAACTCTGCTCACAGCAGGT	33443
Qy	3125	TAAATCTTAGGRCCTAARATTTATCCAAAGGCA	CCARGGCCCTCAGTGAAGGAAYATCCA	3184
Db	33444	-----TACTTAGGGCTAATAATTTCCAAAGGCA	CCAGGGCCCTCAGTGAAGGAACATCCA	33499
Qy	3185	GCCTATCTGGCTTATCCTCATCYCAAAACCCCT	TAAAGCAACTAAGRRRTTCTTGGCCT	3244
Db	33500	GCCTATCTGGCTTATCCTCATCCCAAAACCCCT	TAAAGCAACTAAGGGGATTCCTTGGGCT	33559
Qy	3245	AAVAGGYTTCGCGCAAWATGGATCCCAGAGTWTG	GCRAAATAGCCAGGYCATTTAWATA	3304
Db	33560	AATAGGTTTCTGCCGAAATGGATT--CCCAGGT	TATGGCGAAATATGCCAGGTCATTTAAATA	33618
Qy	3305	CASATAATTAAGGAACTCAGAAAGCCAATACCC	ATTTARTAGATGGAYMCTGAAGYMR	3364
Db	33619	CACATAATTAAGGAACTCAGAAAGCCAATACCC	ATTTARTAGATGGGCAACTGAAGTAG	33678
Qy	3365	AAGTGGCTTTCCAGGCCCCCTTAAAGAAGGCC	CTTAAACCCCAAGYCCCAAGTGTTAAGYTTGCC	3424
Db	33679	AAGTGGCTTTCCAGGCCC-----TAACCCCA	AGCCCGCCAGTGTTAAGYTTGCC	33724
Qy	3425	AACRGGCAAGACTTTTSTTATAYATCA	CAGAAAAAACAAGAAAYAGCTCTRGAGTCC	3484
Db	33725	AACAGGCAAGACTTTTCTTCATATGTCACAG-	AAAAAACAGGAATAGCTCTAGGAGTCC	33783
Qy	3485	TTACACAGRTCCBAGGGAYGAGCTTCGAACCV	RTGGCRYACCTGASTAAGGAAATGATG	3544
Db	33784	TTACACAGATCCBAGGGATGAGCTTCGAACCV	RTGGCACTGACTGAAGGAAATGATG	33843
Qy	3545	TAGTGGCAAAAGGTTGRCYTCATTTGTTTAV	GGGTAGTGGTGGCAGTAGCAGYTKTAGTAT	3604
Db	33844	TAGTGGCAAAAGGTTGACCTCATGTTTACGGG	TAGTGGTGGCAGTAGCAGTCTTAGTAT	33903
Qy	3605	CTGAAGCAGTTAAAAATAACAGGGRAGAGAT	CTTACTGTGTGGACACTCTCATGAAGTGA	3664
Db	33904	CTGAAGCAGTTAAAAATAACAGGGAAGAGAT	CTTACTGTGTGGACACTCTCATGATGTA	33963
Qy	3665	AYRGCACTCATCTGTAAAGGAGACTTGTGGC	TGTFCAGACAACYGVTTTACTTAAATPTC	3724
Db	33964	ATGGCACTACTCATCTGTAAAGGAGACTTGTG	GCTGTFCAGACAACCTGTTTACTTAAATGTC	34023
Qy	3725	AGGCTATTACTTTGAARGGCCAGTGC	GCRACTGTGCACTGTGCAACTCTTAAACCCAG	3784
Db	34024	AGGCTATTACTTTGAARGGCCAGTGC	TCGCACTGTGCAACTCTTAAACCCAG	34083
Qy	3785	YCNCAITTTCTCCAGACAATGAAGAAAGATA	RAAYATAACTGTCTCAACAARTAAITTTCTC	3844
Db	34084	CCACATTTCTTCAGACAATGAAGAAAGATA	RAAYATAACTGTCTCAACAARTAAITTTCTC	34143
Qy	3845	AAACCTATGCCACTCGAGGGGACCTT	VTAGARHTCCYTTGACTGATCCYGACCTTCAAC	3904
Db	34144	AAACCTATGCCACTCGAGGGGACCTT	TAGAGGTTCTCTTGACTGATCCCGACC--TCAAC	34202
Qy	3905	TTGTACTACTGGAAGTTCCTTTGTAGAAAA	AGAGACTTCGAAAAGYGGGTATGCACTG	3964
Db	34203	TTGTATACTGATGGGAAGTTCCTTTGTAGAAA	AGAGACTTCGAAAAGTGGGGTATGCACTG	34262
Qy	3965	GTCAGTGATAATGGAATAYTTGAAAGTAAT	TCCCTCAGTCTCCAGGAACCTAGTGTCTVAGCTR	4024
Db	34263	GTCAGTGATAATGGAATAYTTGAAAGTAAT	TCCCTCAGTCTCCAGGAACCTAGTGTCTVAGCTR	34322
Qy	4025	GCAGAACTAATAGCCVTCAYTTKGGGCAC	TAGAAATTAGGAGAGRAAAAAGGGAATAATA	4084
Db	34323	GCAGAACTAATAGCCCTCACTTGGGCAC	TAGAAATTAGGAGAGAAAAAGGGAATAATA	34382
Qy	4085	TATACAGACTCTRAPATATGCTTACCTAGT	CTCNTCCATGCCCACTGMRGCAATATGSRAGAA	4144
Db	34383	TATACAGACTCTAAATATGCTTTACCTTAGT	CTCTCCATGCCCCTATGCAAGCAATATGGAAGAA	34442
Qy	4145	AGGGAATTCCTAACTTCYAGRGAACACCT	TATCANACATCAGGAAGCCATTAGGARAATTA	4204

34443	AGGGAATTCCTAACTTCTGAGAGAACCTTATCAAAACATTCAGGAAGCCATTAGGAATAATTA	34502
4205	TTATYTGCGCWGTACAGAAACCTTAAGAGAGGTGEMAGTCTTACACTGCVGGGGTTCATCANAAA	4264
34503	TTATTTGGCTGTACAGAAACCTTAAGAGAGGTGGAGTCTTTACACTGCGGGGTTCATCAGAAA	34562
4265	GGAAAGRAAGGGAATAASAGRGAAYTGCCAAAGCAKATATTTGAAGCMAAAGAGCTGCA	4324
34563	GGAAAGGAAGGGAAATAGAAGAGAACTGCGCAAGCAGATATTTGAAGCCAAAAGAGCTGCA	34622
4325	AGGCAGGACCTTCATTTAGAAAATGCTTTATTAACCTTCCTTAGTATATAGGTTAATCCCTTC	4384
34623	AGGCAGGACCTTCATTTAGAAAATGCTTTATTAACCAACCCCTAGTATAGGTTAATCCCTTC	34682
4385	CGGGAAACCAAGCCCGCTAGTACTCAGCAGGAGAAACAGAAATGGGAACTTCACGAGG-CAG	4443
34683	CGGGAAACCAAGCCCGCTAGTACTCAGCAGGAGAAACAGAAATGGGAACTTCACGAGGAGCAG	34742
4444	TTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAAAAATACATTTTGCCTGCGCAACTAT	4503
34743	TTTTCTCCCTCGGACGGCTAGCCACTGAAGAGGGAAAAATACATTTTGCCTGCGCAACTAT	34802
4504	CCAAATGGAATAATTATAAACCTTTCATAAACCTTTCTCTTAGGATCGATAGCACCCA	4563
34803	CCAAATGGAATAATTATAAACCTTTCATAAACCTTTCTCTTAGGATCGATAGCACCCA	34862
4564	TCARATGGCCAAATCATTTATTACTGACCCAGGSCCTTTTCAAACTATCAAGCABATAKT	4623
34863	TCARATGGCCAAATCATTTATTACTGACCCAGGSCCTTTTCAAACTATCAAGCAGATAGT	34922
4624	CAGGCGCTGTGAATGTGCCARAAATAATCCCTGCGCTYATCGCAAGCTCTCTCAGG	4683
34923	CAGGCGCTGTGAGTGTGCCAGAAATAATCCCTGCGCTTATCGCAAGCTCTCTCAGG	34982
4684	APAAACAAARAACAGGCATTACCTGPARAARACTGGCAACTGATTTTACCAAGGCC	4743
34983	AGAAACAAAGAACAGGCCATTACCTGGAGAAGACTGGCAACTGATTTTACCCACAGGCC	35042
4744	AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATACCTTCAACGGTTGGGCARAG	4803
35043	AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTAGATCATTTTCAACGGTTGGGCARAG	35102
4804	GCCTTCCCTGTAGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCATGAATA	4863
35103	GCCTTCCCTGTAGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCATGAATA	35162
4864	ATTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCC	4923
35163	ATTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCC	35222
4924	ACAGTAAACCCAGGAGGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	4983
35223	ACAGTAAACCCAGGAGGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	35282
4984	AGSCCACAAGTCTCAGGGAGAGGTTCGAGAAATCAATGAAGAACTCAAGAGCACTCTAAA	5043
35283	AGSCCACAAGTCTCAGGGAGAGGTTCGAGAAATCAATGAAGAACTCAAGAGCACTCTAAA	35342
5044	AAGCAAAACCCAGGAAACCCACTCACATGGCCCTGYCTGTGTGCTATAGCCTTAAAAAGA	5103
35343	AAGCAAAACCCAGGAAACCCACTCACATGGCCCTGYCTGTGTGCTATAGCCTTAAAAAGA	35402
5104	ATCTGGAATTTCCCCAAAAGAGGACTTAGGCCATACGAAATGCTGTATGGAGGGCC	5163
35403	ATCTGGAATTTCCCCAAAAGAGGACTTAGGCCATACGAAATGCTGTATGGAGGGCC	35462
5164	TTTCAATACCAATGACTTGTGCTTGACCCAGAGACGCCAATTTAGTTGACAGATCACT	5223
35463	TTTCAATACCAATGACTTGTGCTTGACCCAGAGACGCCAATTTAGTTGACAGATCACT	35522
5224	CCTTAGCCAAATATCAACAGTTCTTTAAAAACATTAAGGAACCTTATCCCTCAGAGAGG	5283

Db	35523	CCTTAGCCAAATATCAAAAGTTCTTAAACATTTACAAGGAACCTATCCCTGAGAGAGG	35582
Qy	5284	GAAGAAGCAATTTCCACCCWGGTGCATAGTATTAAGTCAAGTCCTTCYCTCTAAATTC	5343
Db	35583	GAAGAAGCAATTTCCACCCWGGTGCATAGTATTAAGTCAAGTCCTTCYCTCTAAATTC	35642
Qy	5344	CATCCCTAGATACATCTCGGAGAGCCCTACCCAGTCAATTTATYATACCCCACTGCGG	5403
Db	35643	CATCCCTAGATACATCTCGGAGAGCCCTACCCAGTCAATTTATYATACCCCACTGCGG	35702
Qy	5404	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACCTTGGAGTCAAAATCCTGGATCTGC	5463
Db	35703	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACCTTGGAGTCAAAATCCTGGATCTGC	35762
Qy	5464	CAAGAAGCACTGAATTCAGAGGAGCAACCTAGCTATTTCTGTGAACCTCTAGAGATT	5523
Db	35763	CAAGAAGCACTGAATTCAGAGGAGCAACCTAGCTATTTCTGTGAACCTCTAGAGATT	35822
Qy	5524	TGCGCTGCTCTTCAAAACAACACAGAGGAGAAAGTAACTAAATCATAAATCCCCCATG	5583
Db	35823	TGCGCTGCTCTTCAAAACAACACAGAGGAGAAAGTAACTAAATCATAAATCCCCCATG	35881
Qy	5584	GSCTCCCTTATCATATTTTCTCTKTPASTGTTSTTTTAACTTCTTCTCACTCTCACTGCA	5643
Db	35882	GSCTCCCTTATCATATTTTCTCTKTPASTGTTSTTTTAACTTCTTCTCACTCTCACTGCA	35941
Qy	5644	CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGAGAAATG	5703
Db	35942	CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGAGAAATG	36001
Qy	5704	CAGCGTCCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCAAC	5763
Db	36002	CAGCGTCCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCAAC	36061
Qy	5764	TTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCCACTCTTTTGCATGCAT	5823
Db	36062	TTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCCACTCTTTTGCATGCAT	36121
Qy	5824	GCAATATCTATTTTGGACAGAAATAATTAATCTCTAGTTGTCTCGAGGACTTGA	5883
Db	36122	GCAATATCTATTTTGGACAGAAATAATTAATCTCTAGTTGTCTCGAGGACTTGA	36181
Qy	5884	GTCACTGTCTGTGGACTTACTTCCCAACTGTTATGCTGATCGGGGTGAGTTCAA	5943
Db	36182	GTCACTGTCTGTGGACTTACTTCCCAACTGTTATGCTGATCGGGGTGAGTTCAA	36241
Qy	5944	GATCAGGCAAGAGAAAAATGTAAAGAGTAACTCTCCCACTCAACCGGGTACATGGC	6003
Db	36242	GATCAGGCAAGAGAAAAATGTAAAGAGTAACTCTCCCACTCAACCGGGTACATGGC	36301
Qy	6004	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	6063
Db	36302	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	36361
Qy	6064	ACTCGCTGTAGCTTATTAATACACCTCACTGGGCTCCATGAGTCTCGGCCAA	6123
Db	36362	ACTCGCTGTAGCTTATTAATACACCTCACTGGGCTCCATGAGTCTCGGCCAA	36421
Qy	6124	AACCCCTACTAACTGTGGATATGSCCTCCCTGAACTTTCARGCCATATGTTTCAATCCCT	6183
Db	36422	AACCCCTACTAACTGTGGATATGSCCTCCCTGAACTTTCARGCCATATGTTTCAATCCCT	36481
Qy	6184	GTACCTGAAACAATGGAAACAATTTAGCAGCAGAAATAAACAACCTTTTGTAGTGA	6243
Db	36482	GTACCTGAAACAATGGAAACAATTTAGCAGCAGAAATAAACAACCTTTTGTAGTGA	36541
Qy	6244	CCTCTTGTTCATSTGAAATAACCCATACCTCAAACTCACTGTGTAAATTTAGC	6303
Db	36542	CCTCTTGTTCATSTGAAATAACCCATACCTCAAACTCACTGTGTAAATTTAGC	36601
Qy	6304	AATACTACATACACCACTCCCAATGATCAGTGGGTAACTCTCCCAACAATA	6363
Db	36602	AATACTACATACACCACTCCCAATGATCAGTGGGTAACTCTCCCAACAATA	36661
Qy	6364	GTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTTGAAT	6423
Db	36662	GTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTTGAAT	36721
Qy	6424	GGCTCTTTCAGAACTATGTCTTCTCTCATTTCTAGTGGCCCTCATCTACCT	6483
Db	36722	GGCTCTTTCAGAACTATGTCTTCTCTCATTTCTAGTGGCCCTCATCTACCT	36781
Qy	6484	GAACAAGATTTATACAGTTATGTCTATCTAAAGCCCGCAACAAAGAGTACCATTTCT	6543
Db	36782	GAACAAGATTTATACAGTTATGTCTATCTAAAGCCCGCAACAAAGAGTACCATTTCT	36841
Qy	6544	CTTTTGTATAGGACGAGAGTGTAGTGTCACTAGGTACTGGCAATGCGGTATCACCA	6603
Db	36842	CTTTTGTATAGGACGAGAGTGTAGTGTCACTAGGTACTGGCAATGCGGTATCACCA	36901
Qy	6604	ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGGTC	6663
Db	36902	ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGGTC	36961
Qy	6664	GCGACTCCCTGCTGCTACCTTGCAGATCACTTAACTCCCTAGCAGCAGTAGTCTTCRA	6723
Db	36962	GCGACTCCCTGCTGCTACCTTGCAGATCACTTAACTCCCTAGCAGCAGTAGTCTTCRA	37021
Qy	6724	AATCGAAGAGCTTTAGACTYGTCTAACCGCTGAPAGAGGGGAACTGTATTTTAGGG	6783
Db	37022	AATCGAAGAGCTTTAGACTYGTCTAACCGCTGAPAGAGGGGAACTGTATTTTAGGG	37081
Qy	6784	GAAGATGCTGTATTTATGTTTAACTCCGAACTCGTCACTGAGAAAGTTAAAGAAAT	6843
Db	37082	GAAGATGCTGTATTTATGTTTAACTCCGAACTCGTCACTGAGAAAGTTAAAGAAAT	37141
Qy	6844	CSAGATTCGAATCAACCTAGCAGAGAGCTTCGAAACACTGAGCCCTGGGSCCTCCTC	6903
Db	37142	CSAGATTCGAATCAACCTAGCAGAGAGCTTCGAAACACTGAGCCCTGGGSCCTCCTC	37201
Qy	6904	AGCCATGATGATCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTTGCTA	6963
Db	37202	AGCCATGATGATCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTTGCTA	37261
Qy	6964	CTCTCTTTGGACCTCTGATCTTTTACCTCTTTTAACTTTTGTCTCTCCAGAAATCGAA	7023
Db	37262	CTCTCTTTGGACCTCTGATCTTTTAACTCTTTTAACTTTTGTCTCTCTCCAGAAATCGAA	37321
Qy	7024	GCTGTAAACTACTAAATGGAGCCCAAGATGCAAGTCCCAAGTCAAGTCAAGTCAAGTCAAGT	7083
Db	37322	GCTGTAAACTACTAAATGGAGCCCAAGATGCAAGTCCCAAGTCAAGTCAAGTCAAGT	37381
Qy	7084	CTGGACCGGCTGTAGCCACGATCTGATGTTTAAATGATCAATCAAAAGGCAACCTCCTGAG	7143
Db	37382	CTGGACCGGCTGTAGCCACGATCTGATGTTTAAATGATCAATCAAAAGGCAACCTCCTGAG	37441
Qy	7144	GAATCTCAGTGTCAACCTCTACTACGCCCAATTTAGCAGGAGAGTGTAGAGCGT	7203
Db	37442	GAATCTCAGTGTCAACCTCTACTACGCCCAATTTAGCAGGAGAGTGTAGAGCGT	37501
Qy	7204	SGTGGCCCAACCTCCCAACAGCCTTAGGTTTCTGTTGATGGGGGAGCTGAGAGAC	7263
Db	37502	SGTGGCCCAACCTCCCAACAGCCTTAGGTTTCTGTTGATGGGGGAGCTGAGAGAC	37561
Qy	7264	AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAGGTTGACC	7323
Db	37562	AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAGGTTGACC	37621
Qy	7324	ACATCACTTTTAAACAGCGGCTTTGCACTTATGATCACTAGTCACTAGTCACTAGTCACTAGT	7383
Db	37622	ACATCACTTTTAAACAGCGGCTTTGCACTTATGATCACTAGTCACTAGTCACTAGTCACTAGT	37681
Qy	7384	ACTAAATGCTTAATTTAGGCAAGAGAGGAGTAAAGAAATACCAATCACTTATTTGCTG	7443
Db	37682	ACTAAATGCTTAATTTAGGCAAGAGAGGAGTAAAGAAATACCAATCACTTATTTGCTG	37741

```

QY 7444 AGAGCACAGCAGGAGGACATGATCGGATATACCCAGTCTTTCGAGCGCGCAACGG 7503
Db 37742 AGAGCACAGCAGGAGGACATGATCGGATATACCCAGTCTTTCGAGCGCGCAACGG 37801

QY 7504 CAACCCCTTTGGGTCCCTCCCTTGTATGGAGCTCTGTTTTCATGCTATTTTCACTCT 7563
Db 37802 CAACCCCTTTGGGTCCCTCCCTTGTATGGAGCTCTGTTTTCATGCTATTTTCACTCT 37861

QY 7564 ATTAATCTTGCARTGCR 7582
Db 37862 ATTAATCTTGCARTGCA 37880

RESULT 9
LOCUS AC007566/c 149194 bp DNA linear PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.
ACCESSION AC007566
VERSION AC007566.2 GI:11181861
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 149194)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
PUBMED 9847074
REFERENCE 2 (bases 1 to 149194)
AUTHORS Du, Z.
TITLE The sequence of Homo sapiens BAC clone CTB-10G5
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 149194)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 149194)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 149194)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 6 (bases 1 to 149194)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 7 (bases 1 to 149194)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 8 (bases 1 to 149194)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 16, 2000 this sequence version replaced gi:4835815.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc

```

Contact: sapiens@watson.wustl.edu  
 ----- Summary Statistics  
 ----- Center project name: H\_RG010G05  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).  
 VECTOR: pBelobAC11  
 Selection: chloramphenicol

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

The clone CTB-10G5 contains the entire sequence of CTB-83M5.

FEATURES	Location/Qualifiers
source	1..149194 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="7" /map="7q21-7q22" /clone="CTB-10G5" /clone_lib="CTB-978SK-B"
repeat_region	1..1634 /rpt_family="L1"
misc_feature	2248..2388 /note="match to EST BG752883 (NID:gl4063536)"
misc_feature	2248..2387 /note="match to EST AA149693 (NID:gl720635) zn99d06.r1"
misc_feature	2248..2387 /note="match to EST AW579261 (NID:g7254310)"
misc_feature	2248..2387 /note="match to EST BG766882 (NID:gl4077535)"
misc_feature	2248..2387 /note="match to EST BI160365 (NID:gl4620366)"
misc_feature	2248..2374 /note="similar to Homo sapiens EST BF758865 (NID:gl2106765)"
misc_feature	2248..2287 /note="match to EST AV686676 (NID:gl0288539)"
misc_feature	2253..2387 /note="match to EST AU123510 (NID:gl0948226)"
misc_feature	2344..2387







Db	87860	AAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCTTTGACTGATCCCGACC-TCAAC	87802
Qy	3905	TTGTATCTCATGAGAGTTCTTTGTAGAAAAAGGACTTCCAAAAGYGGGGTATGCACTG	3964
Db	87801	TTGTATCTCATGAGAGTTCTTTGTAGAAAAAGGACTTCCAAAAGYGGGGTATGCACTG	87742
Qy	3965	GTCAAGTATATGGAATAYTTGAAAGTAATCCCTCATCTCCAGGAACTAGTGTGCTYAGCTT	4024
Db	87741	GTCAAGTATATGGAATAYTTGAAAGTAATCCCTCATCTCCAGGAACTAGTGTGCTYAGCTT	87682
Qy	4025	GCAGAACTAATAGCCCTTCAYTKGGGCACTAGAAATTTAGGAGAGGAAAGGGAATATA	4084
Db	87681	GCAGAACTAATAGCCCTTCAYTKGGGCACTAGAAATTTAGGAGAGGAAAGGGAATATA	87622
Qy	4085	TATACAGACTCTTARTATGTCTYACCTAGTCTNTCATGTCGCCCATGMRGCAATATGSAAGAA	4144
Db	87621	TATACAGACTCTTARTATGTCTYACCTAGTCTNTCATGTCGCCCATGMRGCAATATGSAAGAA	87562
Qy	4145	AGGGAATTCCTAACTTCYAGAGGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA	4204
Db	87561	AGGGAATTCCTAACTTCYAGAGGAACACCTATCAMACATCAGGAAGCCATTAGGARATTA	87502
Qy	4205	TTATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGYGGGTCTATCANAAA	4264
Db	87501	TTATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGYGGGTCTATCANAAA	87442
Qy	4265	GGAAAGRAAGGGAATASAGGAAAYTGCCAGACAKATATTGAAGCMAAAGAGCTGCA	4324
Db	87441	GGAAAGRAAGGGAATASAGGAAAYTGCCAGACAGATATTGAAGCMAAAGAGCTGCA	87382
Qy	4325	AGGAGGACCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGTAATCCCTTC	4384
Db	87381	AGGAGGACCTCCATTAGAAATGCTTATTAAACTTCCCTTAGTATAGGTAATCCCTTC	87322
Qy	4385	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACCAGG-CAG	4443
Db	87321	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACCAGG-CAG	87262
Qy	4444	TTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGGGAAATACTTTTGCTTGCACACTAT	4503
Db	87261	TTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGGGAAATACTTTTGCTTGCACACTAT	87202
Qy	4504	CMATGGAAATTACTTAAACCCCTTCATCAAACTTTTCACTTAGGATCATAGCACCCA	4563
Db	87201	CMATGGAAATTACTTAAACCCCTTCATCAAACTTTTCACTTAGGATCATAGCACCCA	87142
Qy	4564	TCARATGGCAATCATATTACTGAGCAGGCGCTTTTCAAACTATCAAGCARATAT	4623
Db	87141	TCARATGGCAATCATATTACTGAGCAGGCGCTTTTCAAACTATCAAGCARATAT	87082
Qy	4624	CAGGGCTGTGAATGTGCARAAATAATCCCTGCGCTTATCGCAAGCTCCTTCAGG	4683
Db	87081	CAGGGCTGTGAATGTGCARAAATAATCCCTGCGCTTATCGCAAGCTCCTTCAGG	87022
Qy	4684	ARAACAAARAAACAGGCAATTAACCTGRARAARACTGGCAACTGATTTTACCAAGCCC	4743
Db	87021	ARAACAAARAAACAGGCAATTAACCTGGAGAGAACTGGCAACTGATTTTACCAAGCCC	86962
Qy	4744	AAACCTCAGGATTTTCAGTATCTACTAGTCTGGTARATATTTTACGGGTTGGGCARAG	4803
Db	86961	AAACCTCAGGATTTTCAGTATCTACTAGTCTGGTARATATTTTACGGGTTGGGCARAG	86902
Qy	4804	GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAATA	4863
Db	86901	GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAATA	86842
Qy	4864	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	4923
Db	86841	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	86782
Qy	4924	ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATACATTAACCTGCGCTGCA	4983
Db	86781	ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATACATTAACCTGCGCTGCA	86722
Qy	4984	AGCCCAAGTCTCTCAGGAGAGTTCGAGAAAATGAATGAAYACTCAAAGACATCTAAAA	5043
Db	86721	AGCCCAAGTCTCTCAGGAGAGTTCGAGAAAATGAATGAAYACTCAAAGACATCTAAAA	86662
Qy	5044	AAAGCAACCCAGGAAACCCACCTCACATGGCTGTCTGTGTGCTATAGCTTAAAAAGA	5103
Db	86661	AAAGCAACCCAGGAAACCCACCTCACATGGCTGTCTGTGTGCTATAGCTTAAAAAGA	86602
Qy	5104	ATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	5163
Db	86601	ATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	86542
Qy	5164	TTTCAATACCAATGACTTGTGCTTGACCCCAAGACAGCAACTTTAGTTGACAGATCACCT	5223
Db	86541	TTTCAATACCAATGACTTGTGCTTGACCCCAAGACAGCAACTTTAGTTGACAGATCACCT	86482
Qy	5224	CCTTAGCCAAATATCAACAGTCTTTAAACATTAACAAGAACCTATCCCTGAGAGAGG	5283
Db	86481	CCTTAGCCAAATATCAACAGTCTTTAAACATTAACAAGAACCTATCCCTGAGAGAGG	86422
Qy	5284	GAAGAAGACTATTTCACCCWGTGACATGATATTAGTCAAGTCCCTTCYCTCTAAATCC	5343
Db	86421	GAAGAAGACTATTTCACCCWGTGACATGATATTAGTCAAGTCCCTTCYCTCTAAATCC	86362
Qy	5344	CATCCCTAGATACATCTCGGGAAGGACCTTACCCAGTCAATTTTATYACCCCAACTGCGG	5403
Db	86361	CATCCCTAGATACATCTCGGGAAGGACCTTACCCAGTCAATTTTATYACCCCAACTGCGG	86302
Qy	5404	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATACACTTTGAGTCAAAATCTTGATATGC	5463
Db	86301	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATACACTTTGAGTCAAAATCTTGATATGC	86242
Qy	5464	CAAGGAACCTGAAATCCAGAGACACGCTAGCTATTCTCTGTGACCTCTAGAGGATT	5523
Db	86241	CAAGGAACCTGAAATCCAGAGACACGCTAGCTATTCTCTGTGACCTCTAGAGGATT	86182
Qy	5524	TGCGCTGCTCTTTCAAAACCAACCAAGAGGAAAGTAACCTAAATCATAAATCCCCCATG	5583
Db	86181	TGCGCTGCTCTTTCAAAACCAACCAAGAGGAAAGTAACCTAAATCATAAATCCCCCATG	86123
Qy	5584	GGCTCCCTTATCATATTTTCTCTKATSTGTTSTTTTACCTSTTTTCACTCTCATGCA	5643
Db	86122	GGCTCCCTTATCATATTTTCTCTKATSTGTTSTTTTACCTSTTTTCACTCTCATGCA	86063
Qy	5644	CCCCCTCAGTGGCTGTATGACCACTAGCTCCCTTACCMAGAGTTTCTATGAGAGATT	5703
Db	86062	CCCCCTCAGTGGCTGTATGACCACTAGCTCCCTTACCMAGAGTTTCTATGAGAGATT	86003
Qy	5704	CAGCGTCCCGGAAATATTCATGCTCCCATCTAGGAGTCTTTTAAAGGGAACCCCAACC	5763
Db	86002	CAGCGTCCCGGAAATATTCATGCTCCCATCTAGGAGTCTTTTAAAGGGAACCCCAACC	85943
Qy	5764	TTTCACTGCCCCACACCATATGCCCCGCAACTGCTATCTGCCCCACTCTTTGCACTGCAT	5823
Db	85942	TTTCACTGCCCCACACCATATGCCCCGCAACTGCTATCTGCCCCACTCTTTGCACTGCAT	85883
Qy	5824	GCATACTACTCATTTTGGACAGGAAAAATGATTAATCTAGTTGCTCTGAGGACTTGA	5883
Db	85882	GCATACTACTCATTTTGGACAGGAAAAATGATTAATCTAGTTGCTCTGAGGACTTGA	85823
Qy	5884	GTCACTGTCTGTGGACTTTTACCTCCCAACTGGTATCTCTGATGGGGTGGAGTTCAA	5943
Db	85822	GTCACTGTCTGTGGACTTTTACCTCCCAACTGGTATCTCTGATGGGGTGGAGTTCAA	85763
Qy	5944	GATCAGGCAAGAGAAAAATGTAATAAGAGTAATCTCCCAACTCACCSGGGTACATGGC	6003
Db	85762	GATCAGGCAAGAGAAAAATGTAATAAGAGTAATCTCCCAACTCACCSGGGTACATGGC	85703
Qy	6004	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCAT	6063
Db	85702	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCAT	85643

QY 6064 ACTCGCTGGTAAGCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCAA 6123  
DB 85642 ACTCGCTGGTAAGCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCAA 85593  
QY 6124 AACCTACTAACTGTGGATATGCTCCCTGAACTTCARCCCATATGTTTCAATCCCT 6183  
DB 85582 AACCTACTAACTGTGGATATGCTCCCTGAACTTCAGGCCATATGTTTCAATCCCT 85523  
QY 6184 GTACCTGAACAATGGAACAACCTTCAGCACAGAAATAAACCACACTTCGGTTTACTAGGA 6243  
DB 85522 GTACCTGAACAATGGAACAACCTTCAGCACAGAAATAAACCACACTTCGGTTTACTAGGA 85463  
QY 6244 CCTCTGTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
DB 85462 CCTCTGTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 85403  
QY 6304 AATACTACATACACAACCAACTCCCAATGCATCAGTGGGTAACTCCTCCCAACAATA 6363  
DB 85402 AATACTACATACACAACCAACTCCCAATGCATCAGTGGGTAACTCCTCCCAACAATA 85343  
QY 6364 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATGCTGTTTGAAT 6423  
DB 85342 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATGCTGTTTGAAT 85283  
QY 6424 GCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTAGTGCCTCCCVATGRCATCTACACT 6483  
DB 85282 GCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTAGTGCCTCCCVATGRCATCTACACT 85223  
QY 6484 GAACAAGATTATACAGTTATGCTCATATCTAAGCCCGCAACAAAGAGTAGTACCCATTTCT 6543  
DB 85222 GAACAAGATTATACAGTTATGCTCATATCTAAGCCCGCAACAAAGAGTAGTACCCATTTCT 85163  
QY 6544 CTTTTTGTATAGGAGCAGGAGTGTAGTGCACCTAGGTACTGGCAATGGCGGTATACAA 6603  
DB 85162 CTTTTTGTATAGGAGCAGGAGTGTAGTGCACCTAGGTACTGGCAATGGCGGTATACAA 85103  
QY 6604 ACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGTTC 6663  
DB 85102 ACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGTTC 85043  
QY 6664 GCCGACTCCCTGGTCACCTTGTCAAGATCAACTTAATCCTCTAGCAGCAGTAGTCTCTCRA 6723  
DB 85042 GCCGACTCCCTGGTCACCTTGTCAAGATCAACTTAATCCTCTAGCAGCAGTAGTCTCTCRA 84983  
QY 6724 AATCGAAGAGCTTTAGATCTGCTAATCCGCTGATAGAGGGGAACTGTTTATTTTAGGG 6783  
DB 84982 AATCGAAGAGCTTTAGATCTGCTAATCCGCTGATAGAGGGGAACTGTTTATTTTAGGG 84923  
QY 6784 GAAGATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
DB 84922 GAAGATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 84863  
QY 6844 CSAGATCGAATAACAAGTACAGAGAGCTTCGAAACACCTGGACCTCGGGGCTCTCTC 6903  
DB 84862 CGAGATCGAATAACAAGTACAGAGAGCTTCGAAACACCTGGACCTCGGGGCTCTCTC 84803  
QY 6904 AGCCATGATGATCCCTGGATTTCTCCCTTTTGTAGGAGCTCTAGCAGCTATAATATGCTA 6963  
DB 84802 AGCCATGATGATCCCTGGATTTCTCCCTTTTGTAGGAGCTCTAGCAGCTATAATATGCTA 84743  
QY 6964 CTCCTCTTTGGACCTGTATCTTTACCTCTCTGTTTAACTTTGTTCTCTCCAGAAATCGAA 7023  
DB 84742 CTCCTCTTTGGACCTGTATCTTTAACTCTCTGTTTAACTTTGTTCTCTCCAGAAATCGAA 84683  
QY 7024 GCTGTAAACTACAATGAGGCCCAAGATGCACTCAAGACTTAAGATCTACCGCAGACCC 7083  
DB 84682 GCTGTAAACTACAATGAGGCCCAAGATGCACTCAAGACTTAAGATCTACCGCAGACCC 84623  
QY 7084 CTGGACCGGCTGTTAGCCCAAGATCTGATGTTAATGATCAATCAAGGCCACCTCTCTGAG 7143  
DB 84622 CTGGACCGGCTGTTAGCCCAAGATCTGATGTTAATGATCAATCAAGGCCACCTCTCTGAG 84563  
QY 7144 GAAATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGAGGAGCAGTTAGAGCGGT 7203

DB 84562 GAAATCTCAGCTGCACAACCTCTACTAGCCGCCAATTCAGCAGGAGCAGTTAGAGCGGT 84503  
QY 7204 SGTGGCCAACTCCCCAAACAGACTTAGGTTTCTGTTGAGATGGGGACTGAGAGAC 7263  
DB 84502 CGTGGCCAACTCCCCAAACAGACTTAGGTTTCTGTTGAGATGGGGACTGAGAGAC 84443  
QY 7264 AGGACTAGCTGATTTCTCTAGGCTGATTAAGAAATCCCTAAGCCTAGSTGGGAAGCTGACC 7323  
DB 84442 AGGACTAGCTGATTTCTCTAGGCTGATTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACC 84383  
QY 7324 ACATCCACTTTAAACACGGGCTTGCACCTTAGTTCACCTGACCTGACCAATCAGAGAGCTC 7383  
DB 84382 ACATCCACTTTAAACACGGGCTTGCACCTTAGTTCACCTGACCTGACCAATCAGAGAGCTC 84323  
QY 7384 ACTAAATGCTAATTAGCCAAAGACAGAGGTAAGAAATAGCCCAATCATYATTTGCMTG 7443  
DB 84322 ACTAAATGCTAATTAGCCAAAGACAGAGGTAAGAAATAGCCCAATCATCTATTGCTG 84263  
QY 7444 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGTTCAGCGCGCAACGG 7503  
DB 84262 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGTTCAGCGCGCAACGG 84203  
QY 7504 CAACCCCTTTGGGTCCCTCCCTTTGATCGGAGCTCTGTTTTCATCTATTTCACCTCT 7563  
DB 84202 CAACCCCTTTGGGTCCCTCCCTTTGATCGGAGCTCTGTTTTCATCTATTTCACCTCT 84143  
QY 7564 ATTAATCTTCARCTGCR 7582  
DB 84142 ATTAATCTTCGAATGCA 84124

RESULT 10  
AY101583  
LOCUS  
DEFINITION Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus,  
allele B, complete sequence.  
ACCESSION AY101583  
VERSION AY101583.1 GI:37544401  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 10222)  
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,  
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.  
The endogenous retroviral locus ERVWE1 is a bona fide gene involved  
in hominoid placental physiology  
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)  
14757826  
JOURNAL PUBMED  
REFERENCE 2 (bases 1 to 10222)  
AUTHORS Mallet,F., Bouton,O. and Oriol,G.  
Direct Submission  
TITLE Submitted (06-MAY-2002) Retrovirology Department, UMR 2142  
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée  
d'Italie, Lyon 69364 cedex 07, France  
JOURNAL  
FEATURES  
Location/Qualifiers  
source  
1. 10222  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="21"  
/isolation\_source="PBMC of individual 21"  
/db\_xref="Taxon:9606"  
/chromosome="7"  
/map="7q21-q22"  
/sex="female"  
/cell\_type="PBMC"  
/note="isolated by PCR; ERVWE1 locus, allele B  
endogenous\_virus: HERV-W"  
1. 780  
join(248. .811,2887. .9813)

LTR  
mRNA

```

/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..811,7585..9813)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..811,9248..9813)
/note="putative mRNA transcript 3; no coding region"
7812..9428
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA068162.1"
/db_xref="GI:37544402"
/translation="MALPYHIFLFTVLLPSFTLTAPPCCMTSSPYQBFELMRMORP
GNIDAPYSRLSGKTPFTFTHTHMPNCHYSAILCMHATHYWGKINPSCGGGLV
TVCWYTFQTMGSDGGVQDAKEHAKVEISQLTRMHGTSYKGLDLSKHLTETLR
HTRLVSNLTTLTGLHSEVNTNCWICLPLNFRPYXSVIPVEQWNNEFTSLNTSV
LVGLPVLNLTITNLSLTKVNTTITNCSQICRWTPPTQICLPSGTFPVCVGTSA
YRLCNGSSGMCLVPLPWTITYTEODLSYVVISKPRNKRVPILPVTIGAGVLGALG
TGIGITTSQFYKLSQELNGDMERVADSLVTLQDQLSAAVLQNPALDLLTAE
RGCTCLFLGEECCYYNQSGIVTEKVEIKEDRIQRRABELRNTGPNGLISLQMMWPIILP
FLGDAIAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRLPLDRPAS
RSDVNDIKGTPPBEISAAQPNLPSAGSS"
9487..10222

```

Query Match	86.1%	Score 6446.4	DB 5	Length 10222
Best Local Similarity	95.7%	Pred. No. 0		
Matches 6680	Conservative 195	Mismatches 51	Indels 53	Gaps 27
Qy	606	GATGGGAAACGTTCCCGCAGACAAAAACGCCCTTAAGACGTTATCTGABAAATTTGGGA	665	
Db	2886	GATGGGAAACGTTCCCGCAGACAAAAACGCCCTTAAGACGTTATCTGAGAAATTTGGGA	2945	
Qy	666	MCATTTTGACCCCTCAGACACCTAAAGAAAGAAACGACTTATATCTTCTGCGAGTGCGCCCTG	725	
Db	2946	CCATTTTGACCCCTCAGACACCTAGAAAGAAACGACTTATATCTTCTGCGAGTGCGCCCTG	3005	
Qy	726	GCACCTCTGAGGGAGTATAAATATATAACACCATCTTTACAGCTAGACYTCTTTTGTAGAA	785	
Db	3006	GCACCTCTGAGGGAAGTATAAATATATAACACCATCTTTACAGCTAGACCTCTTTTGTAGAA	3065	
Qy	786	AAGCCAATCGAGTGAAGTGCCATAGTACAAACTTCTTTTCAATTAAGAGACAACTCAC	845	
Db	3066	AAGCCAATCGAGTGAAGTGCCATAGTACAAACTTCTTTTCAATTAAGAGACAACTCAC	3125	
Qy	846	AAATATCTGAAAAAGTGTATTATGCCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC	905	
Db	3126	AAATATCTGAAAAAGTGTATTATGCCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC	3185	
Qy	906	CCAGCATCCCGCACTCTCTCCCAAMTAAATAAGGACCCCCCTTCAACCCAAATGGTCCAA	965	
Db	3186	CCAGCATCCCGCACTCTCTCCCAACTAATAAGGACCCCCCTTCAACCCAAATGGTCCAA	3245	
Qy	966	AAGGAGATGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAATATCCCCAAATATGA	1025	
Db	3246	AAGGAGATGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAATATCCCCAAATATGA	3305	
Qy	1026	CCCTCCAGCACTGGGAGAGAGNAATTCGGGCCAGCAGAGTGCAATGTGCTTTTTYY	1085	
Db	3306	CCCTCTCCAGCACTGGGAGAGAGNAATTCGGGCCAGCAGAGTGCAATGTGCTTTTTTC	3364	
Qy	1086	TCTCCAGACTTAAAGCAAAATAAAACAGACTTTAGGTAAATCTTCAGATAAATCCTCATGG	1145	
Db	3365	TCTCCAGACTTAAAGCAAAATAAAACAGACTTTAGGTAAATCTTCAGATAAATCCTCATGG	3424	
Qy	1146	CTATATTGRTGTTTTTACAAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAGATATA	1205	
Db	3425	CTATATTGATGTTTTTACAAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAGATATA	3483	
Qy	1206	TGTCACCTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCCACCAATACCTGAGCCT	1265	
Db	3484	TGTCACCTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCCACCAATACCTGAGCCT	3543	

Qy	1266	GAGRGTTTGCGGATCTCTGCTATCTCTCAGTCAGGTCAATGGATANGATGACAA	CAGAAGG	1326
Db	3544	GAGAGTTTGCGGATCTCTGCTATCTCTCAGTCAGGTCAATG	-ATAGGATGACAA	CAG-AGG 3600
Qy	1326	AAAGNAATCATTTCCCCACAGGCCAGGCAGTCCTCCAGTCTTASACCCCTCATTTGGGAC	1385	
Db	3601	AAAGAGAAATGATTTCCCCACAGGCCAGGCAGTCCTCCAGTCTAGACCCCTCATTTGGG--	3657	
Qy	1386	ACAGAAATCAGTAACATGGGAGATTTGGTGTGTCGACAGCATTTGCTTAACCTTGCTGCTASAA	1445	
Db	3658	ACACAGAAATCAGAAACAT-GGAGATTTGGTGTGTCGACAGCATTTGCTTAACCTTGCTGCTAGAA	3716	
Qy	1446	GGACTAAGGAAAACTASGAAAGAAARTCTAYGAATTAATCAATGATGTCACACATAACACA	1505	
Db	3717	GGACTAAGGAAAACTAGGAAG-AAGTCTATGAATTAATCAATGATGTCACACATAACACA	3775	
Qy	1506	GGGGAGGGGAAGAAATCCTACTCTGCTTTCTGAGAGACTAAAGGAGGCATTCAGGAAGC	1565	
Db	3776	-GGGAGGGGAAGAAATCCTACTCTGCTTTCTGAGAGACTAAAGGAGGCATTCAGGAAGC	3834	
Qy	1566	GTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTTAATCTTAAAGCGTAACTTTATCA	1625	
Db	3835	GTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTTAATCTTAAAGCGTAACTTTATCA	3894	
Qy	1626	CTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA	1684	
Db	3895	CTCAGTCAGCTGCAGACATTAGAAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA	3954	
Qy	1685	ACTTAGAAACCTTATTGAACCTTGCGAACCTCGGTTTTTTTATTAATAGAGATCAGGAGAGC	1744	
Db	3955	ACTTAGAAACCTTATTGAACCTTGCGAACCTCGGTTTTTTTATTAATAGAGATCAGGAGAGC	4014	
Qy	1745	AGCGGAAACAGGACAAACGGGATTTAAAAAAAAGGCCACCGCTTTTAGTCATGACCCCTCAGG	1804	
Db	4015	AGCGGAAACAGGACAAACGGGATTTAAAAAAAAGGCCACCGCTTTTAGTCATGACCCCTCAGG	4074	
Qy	1805	CAAGTGAGACTTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAATTTGAATGCCTAATAGG	1864	
Db	4075	CAAGTGAGACTTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAATTTGAATGCCTAATAGG	4134	
Qy	1865	GCTTGTCTCCAGTGGCGTCTACAAGACACTTTAAAAAAGATTGTCGAGTAGAAGTAAG	1924	
Db	4135	GCTTGTCTCCAGTGGCGTCTACAAGACACTTTAAAAAAGATTGTCGAGTAGAAGTAAG	4194	
Qy	1925	CCGCCCTCTCGTCCATGCCCTTTATTTCAAGGAAATCAGTGGAAAGCCCACTGCCCCAGG	1984	
Db	4195	CCGCCCTCTCGTCCATGCCCTTTATTTCAAGGAAATCAGTGGAAAGCCCACTGCCCCAGG	4254	
Qy	1985	GGACAAAGGTCTTTTGAGTCAGAGGCACCTAAACAGATATCCAGCAGCAGGACTGAGGG	2044	
Db	4255	GGACAAAGGTCTCTCTGAGTCAGAGGCACCTAAACAGATATCCAGCAGCAGGACTGAGGG	4314	
Qy	2045	TGCTGGGGAAGCGGCATCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTTTGACC	2104	
Db	4315	TGCTGGGGAAGCGGCATCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTTTGACC	4374	
Qy	2105	ATTGAGGGCCAGGAAGGTGTCTCTCCGACACTGGTGGCGTCTTCTTAGTCTTACTCTTC	2164	
Db	4375	ATTGAGGGCCAGG-AGGTGTCTCTCTCCGACACTGGTGGCGTCTTCTTAGTCTTACTCTTC	4433	
Qy	2165	TGTCCTGGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCA	2224	
Db	4434	TGTCCTGGGACAACTGTCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA	4491	
Qy	2225	GTCACTAGATACATTTTATCCAGCCACTAAAGTTATGAATCGGGAGCTTTATTCTTTTCAC	2284	
Db	4492	GTCACTAGATAC-TTCTCCAGCCACTAAAGTTATG-ACTGGGAGCTTTATTCTTTTCAC	4549	
Qy	2285	ATGCTTTTCTAAATATGCTTGAAGCCCACTACCTTGTTAGGAGAGACATTTCTAGCAA	2344	
Db	4550	ATGCTTTTCTAAATATGCTTGAAGCCCACTACCTTGTTAGGAGAGACATTTCTAGCAA	4609	
Qy	2345	AAGCAGGGGCCATTATACACCTTGAAATAGGAGAAAGGAAACACCCGTTGTGTGTGTCCTG	2404	









```
|||||
8955 AATCGAAGAGCTTTAGACTTCTTAACCGCTGAAGAGGGGAACCTGTTTATTTTATGGG 9014
QY GAAGAATGCTGTTATTTATGTTTAAATCCCGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843
Db GAAGAATGCTGTTATTTATGTTTAAATCCCGAATCGTCACTGAGAAAGTTTAAAGAAAT 9074
QY CSAGATCGAATACAGTAKAGCAGAGAGCTTCCGAACACTGGACCTGGGGCTCCTC 6903
Db CGAGATCGAATACAGTAKAGCAGAGAGCTTCCGAACACTGGACCTGGGGCTCCTC 9134
QY AGCCRAATGATCCCTCGAATCTCCCTCTTTAGGACCTCTAGCAGCTATATATTTGCTA 6963
Db AGCCRAATGATCCCTCGAATCTCCCTCTTTAGGACCTCTAGCAGCTATATATTTGCTA 9194
QY CTCCTCTTTGGACCTCTATCTTTTACCTCTCTTTTAACTTTTGTCTCTTCAGAAATCGAA 7023
Db CTCCTCTTTGGACCTCTATCTTTTAACTCTCTTTTAACTTTTGTCTCTTCAGAAATCGAA 9254
QY GCTGTAAACTACAAATGGAGCCCAAGATGCACTGCAAGTCCAAAGCTAAAGTCTACCGCAGACCC 7083
Db GCTGTAAACTACAAATGGAGCCCAAGATGCACTGCAAGTCCAAAGCTAAAGTCTACCGCAGACCC 9314
QY CTGACCGGCTGTTAGCCGATCTGATGTTTAAATGATCAATCAAAAGGCAACCTCTCTGAG 7143
Db CTGACCGGCTGTTAGCCGATCTGATGTTTAAATGATCAATCAAAAGGCAACCTCTCTGAG 9374
QY GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGCAGGAGAGTTCAGAGCGGT 7203
Db GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGCAGGAGAGTTCAGAGCGGT 9434
QY SCTCGGCCAACTCCCAACAGCACTTAGTCTTCTGTTAGATGGGGGACTGAGAGAC 7263
Db SCTCGGCCAACTCCCAACAGCACTTAGTCTTCTGTTAGATGGGGGACTGAGAGAC 9494
QY AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGTGGGAAGGTGACC 7323
Db AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGTGGGAAGGTGACC 9554
QY ACATCCACCTTTAAACAGCGGCTTTGCAACTTTAGTCTCAACCTGACCTAGCTGGGAGGTGACC 7383
Db ACATCCACCTTTAAACAGCGGCTTTGCAACTTTAGTCTCAACCTGACCTAGCTGGGAGGTGACC 9614
QY ACTAAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGT 7443
Db ACTAAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGT 9674
QY AGAGCAGCAGCAGGAGGCAATGATCGGGATATAAACCCAACTTTTCAGCGCGCAACGG 7503
Db AGAGCAGCAGCAGGAGGCAATGATCGGGATATAAACCCAACTTTTCAGCGCGCAACGG 9734
QY CAACCCCTTTGGTCCCTCCCTTTGATGAGGAGCTCTGTTTTTCACTCT 7563
Db CAACCCCTTTGGTCCCTCCCTTTGATGAGGAGCTCTGTTTTTCACTCT 9794
QY ATTAAATCTTGCACTGCR 7582
Db ATTAAATCTTGCACTGCA 9813
```

```
RESULT 11
AY101584
LOCUS Homo sapiens isolate 10222 bp DNA linear. PRI 11-FEB-2004
DEFINITION Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVWE1 locus,
allele A, complete sequence.
ACCESSION AY101584
VERSION AY101584.1 GI:37544403
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
```

```
REFERENCE
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..10222
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="79"
/isolation_source="PBMC of individual 79"
/db_xref="Taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC"
/note="Caucasian"
isolated by PCR; ERVWE1 locus, allele A
endogenous_virus: HERV-W
1..780
join(248..811,2887..9813)
/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..811,7585..9813)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..811,9248..9813)
/note="putative mRNA transcript 3; no coding region"
7812..9428
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68163.1"
/db_xref="GI:37544404"
/translation="MALPYHIFLFTLLPSTLTAPPSCMTSSSPYQEFILWRMORP
TNDAPYVTSLSKGTPTTAHTHMPRNCYHSATLCMHANTHYWGKLNPSCPGILGV
TNCVYTGQMSDGGVQDOARHKVKEVLSOLTVOHGTSSPYKGLDLSKLHETLRT
GIVDAPYVTSLSKGTPTTAHTHMPRNCYHSATLCMHANTHYWGKLNPSCPGILGV
HTRLVSLNTLTGLHVSAGNPNCNCHLPLNRPYVSIIPVEQWNNFSEINTSV
LVGLPVSNLEITHTSNLTCKVPSNTTITNSQCRIRWTPPTQIVCLPSGIFVCGTSA
YRCLNGSESCFLSPFLPMTIYEQDLNYSVSKPRKRPVLPFVIGAGVAGLG
TGIGITSTQFYKLSQELNGDMERVADSLVTLQDLSLAAVVLQNRALDLTAE
RGGTCLFLGECVYVNSQGIIVTEKVEIRDIORRABELRNTGWLSSQWMPWILP
FLGPLAALILLFLPGCIFNLLVNFVSRIRIEAVKLQMEPKMQSKTKIYRRPLDRPAS
PSDNDVNGTPPEISAAQPLLRPNISAGSS"
9487..10222
Query Match 86.0%; Score 6440; DB 5; Length 10222;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6676; Conservative 195; Mismatches 55; Indels 53; Gaps 27;
QY 606 GATGGAAACGTTCCCGCAAGACAAAACGCCCTTAAGACGTAATTCGGAATAATGGGA 665
Db GATGGAAACGTTCCCGCAAGACAAAACGCCCTTAAGACGTAATTCGGAATAATGGGA 2945
QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAAGAGCTTATATCTTCGAGTGGCCCTG 725
Db :
2946 CCAATTTGACCTCAGACACTAAGAAAGAAAGAGCTTATATCTTCGAGTGGCCCTG 3005
QY 726 GCACCTCTGAGGAGAGTATAAATTATACACCATCTTACAGTACACTCTTTCTGTAGAA 785
Db GCACCTCTGAGGAGAGTATAAATTATACACCATCTTACAGTACACTCTTTCTGTAGAA 3065
QY 786 AAGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGACACACTCAC 845
Db AAGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGACACACTCAC 3125
```



Db 5257 ACTTTTGGCGCCCAATTCAGAAACCTTGTGCCATCAGCCACCCCAAGTGCTCTTCAATTT 5316  
QY 3065 CCTCGCYACCTGTGGCTACAMGGTTTCCAAACAPARGCTCARTCTCTCTCACAGCAGGT 3124  
Db 5317 CCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAGGCTCAACTCTCTCTCACAGCAGGT 5376  
QY 3125 TAAATACTTAGGRCTTAARATTTATCCAAAGGCCACGAGGCCCTCAGTGAGGAARYATCCA 3184  
Db 5377 ----TACTTAGGGCTTAAATTTATCCAAAGGCCACGAGGCCCTCAGTGAGGAACACATCCA 5432  
QY 3185 GCCTATACCTGGCTTATCTCTCATCYCAAAACCCCTAAAGCAACTAAGGRBRTTCTTTGGCRT 3244  
Db 5433 GCCTATACCTGGCTTATCTCTCATCCCAAACCCCTAAAGCAACTAAGGGGARTTCTTTGGCGT 5492  
QY 3245 AAYAGGYTCTGCCGAATATGATTTCCCGAGTWTGGCRAAATAGCCAGGYCATPAMATA 3304  
Db 5493 AATAGGTTTCTGCCGAAAATGGATT-CCGAGGTATGGCGAATATAGCCAGGTCAITTAATA 5551  
QY 3305 CASTAATTAAGGAACCTCAGAAAGCCAATACCCATTATTAGATGGAYNCTGAGYMR 3364  
Db 5552 CACTAATTAAGGAACCTCAGAAAGCCAATACCCATTATTAGATGGCAACTGAAGTAG 5611  
QY 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAAGTGTAAAGYTTGCC 3424  
Db 5612 AAGTGGCTTTCCAGGCC-----TAAACCCAGCCCAAGTGTAAAGYTTGCC 5657  
QY 3425 AACRGCGCAAGACTTTTSTYATYRTCAAGAAAAAAACAGRAAYAGCTCTRGAGTCC 3484  
Db 5658 AACGGGCAAGACTTTTCTTCATATGTACAG-AAAAAACAGGAATAGCTCTAGGAGTCC 5716  
QY 3485 TTACACAGTCCRAGGAYGAGCTTGGCAACCVRTGGCRYACCTGASTAAGGAAATGATG 3544  
Db 5717 TTACACAGATCCGAGGGATGAGCTTGGCAACCTGTGGCATACCTGACTAAGGAAATGATG 5776  
QY 3545 TAGTGGCAAGGGTGRCTCAATCTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTAT 3604  
Db 5777 TAGTGGCAAGGGTGTACCTCAITGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT 5836  
QY 3605 CTGAAGCAGTTAAAAATATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
Db 5837 CTGAAGCAGTTAAAAATATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTA 5896  
QY 3665 AVRGCATCTACTGCTAAAGGAGACTTGTGGCTGTCCAGAACYVGTTTACTTAAATRTC 3724  
Db 5897 ATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTCCAGAACCTGTTTACTTAAATGTC 5956  
QY 3725 AGGCTCTATTACTGAAGGCCAGTGTGCRACCTGTGCACTGTGCACTCTTAAACCCAG 3784  
Db 5957 AGGCTCTATTACTGAAGGCCAGTGTGCGACTGTGCACTGTGCACTCTTAAACCCAG 6016  
QY 3785 YCNCAATTTCTCCAGACATGAAGAAAAAGATARAAYATAACTGTGCAACAARTAAATTTCTC 3844  
Db 6017 CCACATTTCTCCAGACATGAAGAAAAAGATAAACAATAACTGTCAACAAGTAAATTTCTC 6076  
QY 3845 AAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCVYTTGACTGATCCYGAACCTTCAAC 3904  
Db 6077 AAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCTCTTGACTGATCCCGACC-TCAAC 6135  
QY 3905 TTGTATACTGATGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATCGATG 3964  
Db 6136 TTGTATACTGATGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGTGGGGTATCGATG 6195  
QY 3965 GTCAGTATAATGGAATAYTTGAAAGTAAATCCCTCACTCCAGGAACCTAGTCTYAGCTR 4024  
Db 6196 GTCAGTATAATGGAATCTTGAAGTAATCCCTCACTCCAGGAACCTAGTCTCAGCTA 6255  
QY 4025 GCAGAACTAATAGCCVTCATYTKGGGCCTAGAAATTAGGAGAGRAAAAAGGGYAAATATA 4084  
Db 6256 GCAGAACTAATAGCCCTCACTTGGGCACCTAGAAATTAGGAGAGAAAAAGGGCAATATA 6315  
QY 4085 TATACAGACTCTRATATCTYACCTAGTCTNTCCATGCCCATGMRGCAATATGSRAGAA 4144

Db 6316 TATACAGACTCTAAATATGCTTACTAGTCTCTCATGCCCATGCGACCAATATGGAAGAA 6375  
QY 4145 AGGAAATTCCTAATCTTCYAGRGAACACCTTATCMACATCAGGAAGCCATTAGGARATTA 4204  
Db 6376 AGGAAATTCCTAATCTTCYAGRGAACACCTTATCMACATCAGGAAGCCATTAGGAATTA 6435  
QY 4205 TTAATGCGWGTACAGAAACCTTARAGAGTGTGMAGTCTTTACACTGCTCYGGGTTCATCANAAA 4264  
Db 6436 TTAATGCGTGTACAGAAACCTTAAAGAGTGGCAGTCTTACACTGCCGGGTTCATCAGAAA 6495  
QY 4265 GGAAGRAAAGGGGAAATASAAAGCAATATGCAACAKATATTTGAAGCMAAAGAGCTGCA 4324  
Db 6496 GGAAGRAAAGGGGAAATAGAGAGAACTGCCAAGCAGATATTTGAAGCCAAAGAGCTGCA 6555  
QY 4325 AGGAGAGACCTTCATTTAGAAATGCTTATTAACCTTCCCTTAGTATATAGGGTAATCCCTTC 4384  
Db 6556 AGGAGAGACCTTCATTTAGAAATGCTTATTAACCAACCCCTTAGTATATAGGGTAATCCCTTC 6615  
QY 4385 CGGGAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCACGAGG-CAG 4443  
Db 6616 CGGGAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCACGAGGACAG 6675  
QY 4444 TTTTCTCCCTCGGGACGGTTAGCCACTGGAAGAGGGGAAAAATACCTTTTTCCTGCAACTAT 4503  
Db 6676 TTTTCTCCCTCGGGACGGTTAGCCACTGGAAGAGGGGAAAAATACCTTTTTCCTGCAACTAT 6735  
QY 4504 CCAATGGAATTAATTTAAACCTTTCATCAAACTTTTCATTTAGGATCGATAGCACCCA 4563  
Db 6736 CCAATGGAATTAATTTAAACCTTTCATCAAACTTTTCATTTAGGATCGATAGCACCCA 6795  
QY 4564 TCARATGCCAAATCAATTTTACTTGGACCAAGGCTTTTCAAAAATCAACAGCAGATAT 4623  
Db 6796 TCARATGCCAAATCAATTTTACTTGGACCAAGGCTTTTCAAAAATCAACAGCAGATAT 6855  
QY 4624 CAGGCTGTGAATGTGCCARARAAATAATTCCTGCTCTATGCCAAGCTCTTTCAGG 4683  
Db 6856 CAGGCTGTGAATGTGCCAGAGAAATAATTCCTGCTCTATGCCAAGCTCTTTCAGG 6915  
QY 4684 ARAACAAARAAACAGGCCAATTAACCTGARAARAACTGGCAACTGATTTTACCCCAAGGCC 4743  
Db 6916 AGAACAAAGAACAGGCCAATTAACCTGGAGAAAGACTGGCAACTGATTTTACCCCAAGGCC 6975  
QY 4744 AACTCAGGATTTTCAGTACTTACTGTGGGTARATACTTTTCCAGGGTGGGCARAG 4803  
Db 6976 AACTCAGGATTTTCAGTACTTACTGTGGGTARATACTTTTCCAGGGTGGGCARAG 7035  
QY 4804 GCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATAATA 4863  
Db 7036 GCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATAATA 7095  
QY 4864 ATTCAGATTCGAGCTTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGGCC 4923  
Db 7096 ATTCAGATTCGAGCTTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGGCC 7155  
QY 4924 ACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCCTGA 4983  
Db 7156 ACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCCTGA 7215  
QY 4984 AGGCCACAGTCTCAGGGAAGGTCGAGAAAATGAATAAAVACTCAAAAGGACATCTAAA 5043  
Db 7216 AGGCCACAGTCTCAGGGAAGGTCGAGAAAATGAATAAAVACTCAAAAGGACATCTAAA 7275  
QY 5044 AAGCAACCCAGGAAACCCACTCACATGGCTGTCTGTGGCTTATAGCTTTAAAGAA 5103  
Db 7276 AAGCAACCCAGGAAACCCACTCACATGGCTGTCTGTGGCTTATAGCTTTAAAGAA 7335  
QY 5104 ATCTCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAATCTGCTATGGAAGGCC 5163  
Db 7336 ATCTCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAATCTGCTATGGAAGGCC 7395  
QY 5164 TTCAATAACCAATGACCTTGTGCTTGACCCCAAGACAGCAACTTGTGTCAGACATCACCT 5223  
Db 7396 TTCAATAACCAATGACCTTGTGCTTGACCCCAAGACAGCAACTTGTGTCAGACATCACCT 7455

QY 5224 CCTTAGCCAAATATCAACAGTCTTAAACATTACAGGAACCTATCCCTGAGAGAGG 5283  
DB |||||  
QY 7456 CCTTAGCCAAATATCAACAGTCTTAAACATTACAGGAACCTATCCCTGAGAGAGG 7515  
DB |||||  
QY 5284 GAAAGAACTATTCCACCCWGTGACATGTTATAGTCAAGTCCCTTCYCTCTAAATCCC 5343  
DB |||||  
QY 7516 GAAAGAACTATTCCACCCCTGTGTGACATGTTATAGTCAAGTCCCTTCCTCTAATCCC 7575  
DB |||||  
QY 5344 CATCCCTAGATACATCCTGGAGAGACCTTACCAGTCAATTTATYACCCCAACTGCGG 5403  
DB |||||  
QY 7576 CATCCCTAGATACATCCTGGAGAGACCTTACCAGTCAATTTATCTACCCCAACTGCGG 7635  
DB |||||  
QY 5404 TTAAGTGGCTGAGTGGAGTCTTTGGATACATCACACTTGAGTCAATCTGATACTGC 5463  
DB |||||  
QY 7636 TTAAGTGGCTGAGTGGAGTCTTTGGATACATCACACTTGAGTCAATCTGATACTGC 7695  
DB |||||  
QY 5464 CAAAGGAACCTGAAATCCAGGAGACAAACCTAGTCTATTCCTGTGAACCTCTAGAGGATT 5523  
DB |||||  
QY 7696 CAAAGGAACCTGAAATCCAGGAGACAAACCTAGTCTATTCCTGTGAACCTCTAGAGGATT 7755  
DB |||||  
QY 5524 TGGCCCTGCTCTTCAAAACAAACAGAGGAGAAAGTAACTAAATCATAAATCCCCCATG 5583  
DB |||||  
QY 7756 TGGCCCTGCTCTTCAAAACAAACAGAGGAGAAAGTAACTAAATCATAAAT - CCCCATG 7814  
DB |||||  
QY 5584 GSCCTCCCTATCATATTTCTCTXTASTGTSTTTTACCCTSTTTCACTCTCACTGCA 5643  
DB |||||  
QY 7815 GSCCTCCCTATCATATTTCTCTXTASTGTSTTTTACCCTSTTTCACTCTCACTGCA 7874  
DB |||||  
QY 5644 CCCCCTCCATGCGCTGTATGACAGTAGTCTCCCTYACCMAGAGTTCTTATGGAGAAAG 5703  
DB |||||  
QY 7875 CCCCCTCCATGCGCTGTATGACAGTAGTCTCCCTYACCMAGAGTTCTTATGGAGAAAG 7934  
DB |||||  
QY 5704 CAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACC 5763  
DB |||||  
QY 7935 CAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACC 7994  
DB |||||  
QY 5764 TTCACCTGCCACACCCATGATGCCCGCAACTGCTATCACTGCGCACTCTTTGCAATGCA 5823  
DB |||||  
QY 7995 TTCACCTGCCACACCCATGATGCCCGCAACTGCTATCACTGCGCACTCTTTGCAATGCA 8054  
DB |||||  
QY 5824 GCAAACTACTCATTTATGGACAGGAAATGATTAACTCTAGTGTCTCGGAGGACTTGA 5883  
DB |||||  
QY 8055 GCAAACTACTCATTTATGGACAGGAAATGATTAACTCTAGTGTCTCGGAGGACTTGA 8114  
DB |||||  
QY 5884 GTCACCTGTCTGTGACTTACCCAACTGTATGCTCTCATGCGGGGTGAGTTCAA 5943  
DB |||||  
QY 8115 GTCACCTGTCTGTGACTTACCCAACTGTATGCTCTCATGCGGGGTGAGTTCAA 8174  
DB |||||  
QY 5944 GATCAGGCAAGAGAAACATGTAAAGAGTAACTCTCCAACTCAACCGGGGTACATGGC 6003  
DB |||||  
QY 8175 GATCAGGCAAGAGAAACATGTAAAGAGTAACTCTCCAACTCAACCGGGGTACATGGC 8234  
DB |||||  
QY 6004 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCGTACCCAT 6063  
DB |||||  
QY 8235 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCGTACCCAT 8294  
DB |||||  
QY 6064 ACTCGCCTGTAAGCTATTTAATACACCCCTCACTGGGCTCCATGAGTCTCGGCCCAA 6123  
DB |||||  
QY 8295 ACTCGCCTGTAAGCTATTTAATACACCCCTCACTGGGCTCCATGAGTCTCGGCCCAA 8354  
DB |||||  
QY 6124 AACCTCTACTAACTGTGTGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 6183  
DB |||||  
QY 8355 AACCTCTACTAACTGTGTGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 8414  
DB |||||  
QY 6184 GTACCTGAACATGGAACAACTTTCAGCAGAGAAATAAACAACAACCTTCGTTTATGAGA 6243  
DB |||||  
QY 8415 GTACCTGAACATGGAACAACTTTCAGCAGAGAAATAAACAACAACCTTCGTTTATGAGA 8474  
DB |||||  
QY 6244 CCTCTGTTTCCAAATGGAATAACCCATACCTCAAACTCACCTGTGTAATAATTTAGC 6303  
DB |||||  
QY 8475 CCTCTGTTTCCAAATGGAATAACCCATACCTCAAACTCACCTGTGTAATAATTTAGC 8534  
DB |||||

QY 6304 AATACTACATACACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCTCCACACAAATA 6363  
DB |||||  
QY 8535 AATACTACATACACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCTCCACACAAATA 8594  
DB |||||  
QY 6364 GTCTGCTTACCTCAGGAATATTTTGTGTGTGTACCTCAGCCTATCGTCTTGTGAAT 6423  
DB |||||  
QY 8595 GTCTGCTTACCTCAGGAATATTTTGTGTGTGTGTACCTCAGCCTATCGTCTTGTGAAT 8654  
DB |||||  
QY 6424 GGTCTTTCAGAACTATAGTGTCTCTCTCATTTAGTGCCTCCCTATGRCCTATCACT 6483  
DB |||||  
QY 8655 GGTCTTTCAGAACTATAGTGTCTCTCTCATTTAGTGCCTCCCTATGACCACTACACT 8714  
DB |||||  
QY 6484 GACAGAGATTTATACAGTTATGTCTATCTAAGCCCGCCAAACAAAGAGTACCCATCTT 6543  
DB |||||  
QY 8715 GACAGAGATTTATACAAATTTATGTCTATCTAAGCCCGCCAAACAAAGAGTACCCATCTT 8774  
DB |||||  
QY 6544 CTTTTTGTATAGGAGCAGAGTGTCTAGTGTGCACTAGGTACTGGCAATGCGGTATCACA 6603  
DB |||||  
QY 8775 CTTTTTGTATAGGAGCAGAGTGTCTAGTGTGCACTAGGTACTGGCAATGCGGTATCACA 8834  
DB |||||  
QY 6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 6663  
DB |||||  
QY 8835 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 8894  
DB |||||  
QY 6664 GCGGACTCCCTGCTGACCTTGCAAGATCAACTTTAATCTCCTAGCAGCAGTAGTCTTCRA 6723  
DB |||||  
QY 8895 GCGGACTCCCTGCTGACCTTGCAAGATCAACTTTAATCTCCTAGCAGCAGTAGTCTTCRA 8954  
DB |||||  
QY 6724 AATCGAAGAGCTTTAGACTGTCTAACCGCTGAPAGAGGGGACCTGTTTATTTAGGG 6783  
DB |||||  
QY 8955 AATCGAAGAGCTTTAGACTGTCTAACCGCTGAPAGAGGGGACCTGTTTATTTAGGG 9014  
DB |||||  
QY 6784 GAAGAACTGTGTTTATTTATGTTAAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
DB |||||  
QY 9015 GAAGAACTGTGTTTATTTATGTTAAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 9074  
DB |||||  
QY 6844 CSAGATTCGAATACAACTGATAGCAGAGACTTCGAAACACTTGGACCTTGGGSCCTCCTC 6903  
DB |||||  
QY 9075 CGAGATTCGAATACAACTGATAGCAGAGAGCTTCGAAACACTTGGACCTTGGGSCCTCCTC 9134  
DB |||||  
QY 6904 AGCCATGAGTCCCTGGAATCTCCCTTCTTTAGGACCTCTAGCAGCTATATTTGCTA 6963  
DB |||||  
QY 9135 AGCCATGAGTCCCTGGAATCTCCCTTCTTTAGGACCTCTAGCAGCTATATTTGCTA 9194  
DB |||||  
QY 6964 CTCTCTTTGGACCTCTGATCTTTTACCTCTCTTGTAACTTTGTCTCTTCCAGAAATCGAA 7023  
DB |||||  
QY 9195 CTCTCTTTGGACCTCTGATCTTTTAACTCTCTTGTAACTTTGTCTCTTCCAGAAATCGAA 9254  
DB |||||  
QY 7024 GCTGTAAACTACAAATGGAGCCCAAGATGCACTCCAGACTAAGATCTACCGCAGACCC 7083  
DB |||||  
QY 9255 GCTGTAAACTACAAATGGAGCCCAAGATGCACTCCAGACTAAGATCTACCGCAGACCC 9314  
DB |||||  
QY 7084 CTGGACCGGCTGTAGCCCAAGATGCACTCCAGACTAAGATCTACCGCAGACCC 9374  
DB |||||  
QY 9315 CTGGACCGGCTGTAGCCCAAGATGCACTCCAGACTAAGATCTACCGCAGACCC 9374  
DB |||||  
QY 7144 GAAATCTCAGCTGCAACAACTCTACTCGCCCAATTCAGCAGGAGCAGTTAGAGCGGT 7203  
DB |||||  
QY 9375 GAAATCTCAGCTGCAACAACTCTACTCGCCCAATTCAGCAGGAGCAGTTAGAGCGGT 9434  
DB |||||  
QY 7204 GGTGGCCCAACCTCCCAACAGCACTTAGGTTTCTCTGTGAGATGGGGGACTGAGAGAC 7263  
DB |||||  
QY 9435 CATCGGCCAACTCCCAACAGCACTTAGGTTTCTCTGTTGAGATGGGGGACTGAGAGAC 9494  
DB |||||  
QY 7264 AGGACTAGCTGATTTCTTAGCTGATTAAGAACTCCYTAAGCCTAGSTGGGAAGTGACC 7323  
DB |||||  
QY 9495 AGGACTAGCTGATTTCTTAGCTGATTAAGAACTCCYTAAGCCTAGSTGGGAAGTGACC 9554  
DB |||||  
QY 7324 ACATCCACTTTAACAACGGGCTTGCACTTAGTGTCACTGACCTAGCAGAGCTC 7383  
DB |||||  
QY 9555 ACATCCACTTTAACAACGGGCTTGCACTTAGTGTCACTGACCTAGCAGAGCTC 9614  
DB |||||  
QY 7384 ACTAAATGCTAAATTAGGCAAGAGCAGGAGGTAAAGAAATAGCCCAATCATYTTATTGCTG 7443  
DB |||||

```
||||| 9615 ACTAAATGCTAAATTAGCCAAACACAGAGGTGAAGAAATAGCCAAATCATCTATTGCTG 9674
||||| 7444 AGAGCACAGCAGGAGGGAACAATGATCGGGATATAAAACCCAAAGTTCAGCGCGGCAACGG 7503
||||| 9675 AGAGCACAGCAGGAGGGAACAATGATCGGGATATAAAACCCAAAGTTCAGCGCGGCAACGG 9734
||||| 7504 CAACCCCTTTGGGTCCCTCCCTTTGATCGGAGCTCTGTTTCATGCTATTTCACCTCT 7563
||||| 9735 CAACCCCTTTGGGTCCCTCCCTTTGATCGGAGCTCTGTTTCATGCTATTTCACCTCT 9794
||||| 7564 ATTAATCTTGCACTGCR 7582
||||| 9795 ATTAATCTTGCACTGCA 9813

RESULT 12
BD221808 10499 bp DNA linear PAT 17-JUL-2003
LOCUS BD221808
DEFINITION Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses.
ACCESSION BD221808
VERSION BD221808.1 GI:33031578
KEYWORDS JP 2002518051-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 10499)
AUTHORS Alliel,P.M., Perin,J.P. and Rieger,F.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: JP 2002518051-A 3 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Homo sapiens (human)
PN JP 2002518051-A/3
PD 25-JUN-2002
PF 23-JUN-1999 JP 2000556036
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
C12N15/09, A01K67/027, A61K31/711, A61K39/21, A61K48/00, A61P21/00, PC
A61P25/00
PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC Nucleic
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1..10499
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..10499
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 85.9%; Score 6436; DB 2; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

QY 606 GATGGAAACGTTCCCGCAAGACAAAACCGCCCTTAAGACGTATTTCGGAPAAATTGGGA 665
DB 2953 GATGGAAACGTTCCCGCAAGACAAAACCGCCCTTAAGACGTATTTCGGAGAAATTGGGA 3012
QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAAGACGTATTATCTCTGCGTCCGCTG 725
DB 3013 CCAATTTGACCTCAGACACTAAGAAAGAAAGACGTATTATCTCTGCGTCCGCTG 3072
QY 726 GCACCTCCTGAGGGAAGTAAATTAATTAACACCATCTTACAGCTAGACACTTTTGTAGAA 785
DB 3073 GCACCTCCTGAGGGAAGTAAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAA 3132
```

```
QY 786 AAGCAAAATGAGTGAAGTGCATTAAGTACAAAACCTTTCTTTTCAATTAGAGACAACCTCAC 845
DB 3133 AAGCAAAATGAGTGAAGTGCATTAAGTACAAAACCTTTCTTTTCAATTAGAGACAACCTCAC 3192
QY 846 AATTATGTAAAAAGTGTGATTTATGCGCTACAGGAAGCTTTCAGAGTCTACCTCCCTATC 905
DB 3193 AATTATGTAAAAAGTGTGATTTATGCGCTACAGGAAGCTTTCAGAGTCTACCTCCCTATC 3252
QY 906 CCAGATCCCGACTCTCTCCCAATTAATTAAGAGACCCCTTCAACCAATGCTGTCACAA 965
DB 3253 CCAGATCCCGACTCTCTCCCAATTAATTAAGAGACCCCTTCAACCAATGCTGTCACAA 3312
QY 966 AAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCATATTATCCCAATATATGA 1025
DB 3313 AAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCATATTATCCCAATATATGA 3372
QY 1026 CCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTYY 1085
DB 3373 CCCCT-CCAAGCAGTGGGAGGAAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTTC 3431
QY 1086 TCTCCAGACTTAAGCAAAATAAACAAGACTTAGTAAATTTCTCAGATAATTCCTGATGG 1145
DB 3432 TCTCCAGACTTAAGCAAAATAAACAAGACTTAGTAAATTTCTCAGATAATTCCTGATGG 3491
QY 1146 CTATATTGRTGTTTACAAAGGTTAGGACAATTTCTTGATCTGCATCGAGAGATATATA 1205
DB 3492 CTATATTGATGTTTACAAAGGTTAGGACAATTTCTTGATCTGCATCGAGAGATATA-A 3550
QY 1206 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCACCAATAACTGCAGCT 1265
DB 3551 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCACCAATAACTGCAGCT 3610
QY 1266 GAGGTGTCGATCTCTGATCTCAGTCAAGTCAATGATANGATGACACAGAGG 1325
DB 3611 GAGGTGTCGATCTCTGATCTCAGTCAAGTCAATG--ATAGGATGACAAACAG-AGG 3667
QY 1326 AAAGANAATGATCCCAAGCAGCAGCAGTTCACAGTCCACCTCATTTGGGAC 1385
DB 3668 AAAGANAATGATCCCAAGCAGCAGCAGTTCACAGTCCACCTCATTTGGG--- 3724
QY 1386 ACAGAAATCAGTAACATGGAGATTTGTCGACAGACATTTGCTAACTGTGTGTASAA 1445
DB 3725 ACACAGAATCAGAACAT--GGAGATTTGTCGACAGACATTTGCTAACTGTGTGTAGAA 3783
QY 1446 GGACTAAGAAAACTASGAAGAAATCTAYGAATTAATCAATGATGTCACCAATACACA 1505
DB 3784 GGACTAAGAAAACTTAGGAAG--AAGTCTATGAATTTACTCAATGATGTCACCAATACACA 3842
QY 1506 GGGGAAGGAAGAAAAATCCTACTGCTTCTGAGAGACTAAGGAGGAGCATTTGAGAGAGC 1565
DB 3843 -GGGAGGAAGAAAAATCCTACTGCTTCTGAGAGACTAAGGAGGAGCATTTGAGAGAGC 3901
QY 1566 GTGCTCTCTGTCACCTGACTCTTCTGAAAGCCAACTAATCTTAAGCGTAAATTTATCA 1625
DB 3902 GTGCTCTCTGTCACCTGACTCTTCTGAAAGCCAACTAATCTTAAGCGTAAATTTATCA 3961
QY 1626 CTCAGTCAAGTGCAGACATTAG-AAAAAATTTCAAAAGTCTGCCCTAGCCCGGAGCAAA 1684
DB 3962 CTCAGTCAAGTGCAGACATTAGAAAAAATTTCAAAAGTCTGCCCTAGCCCGGAGCAAA 4021
QY 1685 ACTTAGAAACCTTATTTGAACCTTTGGCAACVTCGGTTTTTTTATTAATAGAGATCAGGAGGAGC 1744
DB 4022 ACTTAGAAACCTTATTTGAACCTTTGGCAACVTCGGTTTTTTTATTAATAGAGATCAGGAGGAGC 4081
QY 1745 AGCGGAAACAGGACAAAACGGGATTAATAAAGGACCGCTTTTAGTTCATGACCTTCAGG 1804
DB 4082 AGCGGAAACAGGACAAAACGGGATTAATAAAGGACCGCTTTTAGTTCATGACCTTCAGG 4141
QY 1805 CAAATGAGCTTTGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTAATGCTTAATAGG 1864
DB 4142 CAAATGAGCTTTGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTAATGCTTAATAGG 4201
QY 1865 GCTTCTTCCAGTGGGCTTACAGGACACTTTTAAAAAAGATTTGTCCAAGTAGAAGTAAG 1924
```





QY 4085 TATACAGACTCTBARTATGCTYACCTAGTCTNTCCATGCCCATGMRGCAATATGSRAGAA 4144  
Db 6383 TATACAGACTCTAATATATGCTTTACCTAGTCTCTCCATGCCCATGCGCAATATGGAAGAA 6442  
QY 4145 AGGGAATTCCTAACTTCYAGRGAAACACCTATCAMACATCAGAAAGCCATTAGGARATTA 4204  
Db 6443 AGGGAATTCCTAACTTCYAGRGAAACACCTATCAMACATCAGAAAGCCATTAGGARATTA 6502  
QY 4205 TPATYGGCWGTTACAGAAACCTTARAGAGTGGMAGTCTTACACTGCTCYGGGGTCATCANAAA 4264  
Db 6503 TTTATGGCTGTACAGAAACCTTAAAGAGTGGCAGTCTTACACTGCTCGGGGTTCATCAGAAA 6562  
QY 4265 GGAAAGRAAGGAAATASAGRGAAATGCGCAAGCAKATATTTGAAGCMAAAGAGCTGCA 4324  
Db 6563 GGAAAGRAAGGAAATAGAGAGAACTGCCAAGCAGATATTTGAAGGCCAAAGAGCTGCA 6622  
QY 4325 AGGCAGGACCTCCATTTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC 4384  
Db 6623 AGGCAGGACCTCCATTTAGAAATGCTTATTAACCAACCCCTTAGTATAGGGTAATCCCTTC 6682  
QY 4385 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG-CAG 4443  
Db 6683 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGGACAG 6742  
QY 4444 TTTTCTCCCTCGGAGCGGTTAGCCACTGAAGAGGAAATACCTTTGCTGCGCAACTAT 4503  
Db 6743 TTTTCTCCCTCGGAGCGGTTAGCCACTGAAGAGGAAATACCTTTGCTGCGCAACTAT 6802  
QY 4504 CCAATGGAAATTTACTTAAACCTTCAACAACTTTCATCTAGGCATCGATAGCACCA 4563  
Db 6803 CCAATGGAAATTTACTTAAACCTTCAACAACTTTCATCTAGGCATCGATAGCACCA 6862  
QY 4564 TCARATGGCAATCATTTACTGGACAGGCGCTTTTCAAACTATCAAGCARATAT 4623  
Db 6863 TCARATGGCAATCATTTACTGGACAGGCGCTTTTCAAACTATCAAGCARATAGT 6922  
QY 4624 CAGGCGCTGTGAATGTGCGARARAAATATCCCTGCGCTYATGCGCAAGCTCCTTCAGG 4683  
Db 6923 CAGGCGCTGTGAATGTGCGARARAAATATCCCTGCGCTTATCGCAAGCTCCTTCAGG 6982  
QY 4684 ARAACAAARAAACAGGCCATTCCTGRABARACTGGCAACTGATTTACCCACAGGCC 4743  
Db 6983 AGAACAAAGAACAGGCCATTCCTGGAGAACCTGGCAACTGATTTACCCACAGGCC 7042  
QY 4744 AAACCTCAGGAGTTTCAGTATCTACTAGTCTGGGTARATACCTTTCACGGGTTGGCARAG 4803  
Db 7043 AAACCTCAGGAGTTTCAGTATCTACTAGTCTGGGTARATACCTTTCACGGGTTGGCARAG 7102  
QY 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAATA 4863  
Db 7103 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAATA 7162  
QY 4864 ATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCCAGGCC 4923  
Db 7163 ATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCCAGGCC 7222  
QY 4924 ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA 4983  
Db 7223 ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA 7282  
QY 4984 AGGCCACAGTCTCTCAGGAGGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA 5043  
Db 7283 AGGCCACAGTCTCTCAGGAGGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA 7342  
QY 5044 RAGCAACCCAGGAACCCACCTCACATGCGCTGTCTGTGCTATAGCCCTTAAAGAA 5103  
Db 7343 RAGCAACCCAGGAACCCACCTCACATGCGCTGTCTGTGCTATAGCCCTTAAAGAA 7402  
QY 5104 ATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 5163  
Db 7403 ATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 7462

QY 5164 TTTCAATACCAATGACCTTGTGCTTGACCCAAAGACAGCAACTTAGTTGCGACATCACCT 5223  
Db 7463 TTTCAATACCAATGACCTTGTGCTTGACCCAAAGACAGCAACTTAGTTGCGACATCACCT 7522  
QY 5224 CTTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAGGAACCTTATCCCTGAGAAGAGG 5283  
Db 7523 CTTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAGGAACCTTATCCCTGAGAAGAGG 7582  
QY 5284 GAAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATTTCC 5343  
Db 7583 GAAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATTTCC 7642  
QY 5344 CATCCCTAGATACATCTTGGGAAGAGCCCTACCAAGTCAATTTATATATATATATATAT 5403  
Db 7643 CATCCCTAGATACATCTTGGGAAGAGCCCTACCAAGTCAATTTATATATATATATATAT 7702  
QY 5404 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCAACATTTGAGTCAAAATCTTGATATCTGC 5463  
Db 7703 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCAACATTTGAGTCAAAATCTTGATATCTGC 7762  
QY 5464 CAAAGGAACCTGGAATATCCAGGAGCAACGCTAGCTATTTCTGTGAAACCTCTAGAGGATT 5523  
Db 7763 CAAAGGAACCTGGAATATCCAGGAGCAACGCTAGCTATTTCTGTGAAACCTCTAGAGGATT 7822  
QY 5524 TGGCGCTCTCTTCAACAAACCAACAGGAGGAAGTAATAAATCATAAAT-CCCCATG 5583  
Db 7823 TGGCGCTCTCTTCAACAAACCAACAGGAGGAAGTAATAAATCATAAAT-CCCCATG 7881  
QY 5584 GSCCTCCCTTATCATATTTTCTCTKTAATGTTSTTTTACCCTSTTTTCACTCTCACTGCA 5643  
Db 7882 GSCCTCCCTTATCATATTTTCTCTTACTTGTCTTTTACCTCTTCACTCTCACTGCA 7941  
QY 5644 CCCCCTCATGCGCTGTATGACAGTAGTCTCCCTYACCMAGAGTTTCTATGGAGATG 5703  
Db 7942 CCCCCTCATGCGCTGTATGACAGTAGTCTCCCTTACCAAGAGTTTCTATGGAGATG 8001  
QY 5704 CAGCGTCCGGAATATTTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCCCAAC 5763  
Db 8002 CAGCGTCCGGAATATTTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCCCAAC 8061  
QY 5764 TTCAGTCCCAACACCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCA 5823  
Db 8062 TTCAGTCCCAACACCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCA 8121  
QY 5824 GCAATACTCATTTATGGACAGGAAAAATGATTAATCTCTAGTTGCTGGAGAGCTTGA 5883  
Db 8122 GCAATACTCATTTATGGACAGGAAAAATGATTAATCTCTAGTTGCTGGAGAGCTTGA 8181  
QY 5884 GTCAGTGTCTGTGGACTTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTTCAA 5943  
Db 8182 GTCAGTGTCTGTGGACTTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTTCAA 8241  
QY 5944 GATCAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCGGGGTACATGCG 6003  
Db 8242 GATCAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCGGGGTACATGCG 8301  
QY 6004 ACCTTAGCCCTACAAAGGACTAGATCTCTCAAAATACATAGAAACCCCTCGTACCCAT 6063  
Db 8302 ACCTTAGCCCTACAAAGGACTAGATCTCTCAAAATACATAGAAACCCCTCGTACCCAT 8361  
QY 6064 ACTCGCTGTGAAGCTATTTAATACCACTTACCTGCGCTCCATGAGGTCTCGGCCAA 6123  
Db 8362 ACTCGCTGTGAAGCTATTTAATACCACTTACCTGCGCTCCATGAGGTCTCGGCCAA 8421  
QY 6124 AACCCCTACTACTGTGTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCCCT 6183  
Db 8422 AACCCCTACTACTGTGTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCCCT 8481  
QY 6184 GTACCTGAACAAATGGAACAACTTCAGCAAGAAAAATAACACCACTTCCGTTTTAGTAGGA 6243  
Db 8482 GTACCTGAACAAATGGAACAACTTCAGCAAGAAAAATAACACCACTTCCGTTTTAGTAGGA 8541  
QY 6244 CCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAACCTCCTGCTGTGAAAAATTTAGC 6303



		:	
8542	Db	CCTCTGTGTTTCCAATCTGGAAATAACCCATACCTCAAACTCACCCTGTGTAAAAATTAGC	8601
6304	Qy	AATACTACATACACAACCAACTCCCAAATGCATCAGSTGGGTAACTCTCTCCACACAATAA	6363
8602	Db	AATACTACATACACAACCAACTCCCAAATGCATCAGSTGGGTAACTCTCTCCACACAATA	8661
6364	Qy	GTCCTGCCCTACCCCTCAGGAATATTTTTTGTCGTGGGTACTCAGCGTAFCGTGTGTGAAT	6423
8662	Db	GTCCTGCCCTACCCCTCAGGAATATTTTTTGTCGTGGGTACTCAGCGTAFCGTGTGTGAAT	8721
6424	Qy	GGCTCTTCAGAAATCTATGTGCTCTCCTCTCATTTCTTAGTGGCCCCCVATGRCCATCTCAC	6483
8722	Db	GGCTCTTCAGAAATCTATGTGCTCTCCTCTCATTTCTTAGTGGCCCCCVATGRCCATCTCAC	8781
6484	Qy	GAAACAAGATTTPATACAGTTATGTGCATATCTAAGCCCCGCCAACAAAAGAGTACCCATTCTT	6543
8782	Db	GAAACAAGATTTPATACAGTTATGTGCATATCTAAGCCCCGCCAACAAAAGAGTACCCATTCTT	8841
6544	Qy	CCITTTTGTTATAGAGACAGAGTGCTTAGTGTGCATAGGTACTGGCATTTGGCGGTATCAC	6603
8842	Db	CCITTTTGTTATAGAGACAGAGTGCTTAGTGTGCATAGGTACTGGCATTTGGCGGTATCAC	8901
6604	Qy	ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTTAAATCGGGACATGGACCGGTC	6663
8902	Db	ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTTAAATCGGGACATGGACCGGTC	8961
6664	Qy	GCCGACTCCCTGGTGCACCTTGCAAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTCRA	6723
8962	Db	GCCGACTCCCTGGTGCACCTTGCAAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTCRA	9021
6724	Qy	AATCGAAGAGCTTTAGACTVGTCTAAACCGCTGARAGAGGGGAACTGTTTATTTTAGGG	6783
9022	Db	AATCGAAGAGCTTTAGACTVGTCTAAACCGCTGARAGAGGGGAACTGTTTATTTTAGGG	9081
6784	Qy	GAAGAATGCTGTATATGTATATCAATCCGGAATCGTCACTGAGAAGTTTAAAGAAATT	6843
9082	Db	GAAGAATGCTGTATATGTATATCAATCCGGAATCGTCACTGAGAAGTTTAAAGAAATT	9141
6844	Qy	CSAGATCGAATACAAACGTAVAGACAGARGAGCTTGAAACACTGSAACCTGGGGCTCCTC	6903
9142	Db	CGAGATCGAATACAAACGTAVAGACAGARGAGCTTGAAACACTGSAACCTGGGGCTCCTC	9201
6904	Qy	AGCCRATGGATGGCCCTGGATTTCTCCCTCTTAGGACCTCTAGCAGCTATAATTTGGCTA	6963
9202	Db	AGCCAATGGATGGCCCTGGATTTCTCCCTCTTAGGACCTCTAGCAGCTATAATTTGGCTA	9261
6964	Qy	CTCCTCTTTGGACCCCTGTATCTTTTRACCTCGTTGTTHAACHTTGTCTCTCCAGAATCGAA	7023
9262	Db	CTCCTCTTTGGACCCCTGTATCTTTTAACTCTCTTTTAACTTTGTCTCTCCAGAATCGAA	9321
7024	Qy	GCTGTAAAACCTACAAATGGAGCCCAAGATCAGTCAGTCCAAAGACTAAGATCTACCCGACACC	7083
9322	Db	GCTGTAAAACCTACAAATGGAGCCCAAGATCAGTCAGTCCAAAGACTAAGATCTACCCGACACC	9381
7084	Qy	CTGACCGGCTGYTAGCCCAACGATCTGTATGTAAATGACATCAAAGACACCCCTCTGTAG	7143
9382	Db	CTGACCGGCTGYTAGCCCAACGATCTGTATGTAAATGACATCAAAGACACCCCTCTGTAG	9441
7144	Qy	GAATCTCAGCTGCAACCTCTACTAGGCCCAATTCACAGGAAGCAGTTAGACGGGT	7203
9442	Db	GAATCTCAGCTGCAACCTCTACTAGGCCCAATTCACAGGAAGCAGTTAGACGGGT	9501
7204	Qy	SGTCGGCCACCTCCCAACAGCACTTAGTTTTCTTTCTGTTGAGATGGGGACATGAGAC	7263
9502	Db	- C - TCGGCCAACCTCCCAACAGCACTTAGTTTTCTTTCTGTTGAGATGGGGACATGAGAC	9560
7264	Qy	AGGACTAGCTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCTTAGSTGGGAAGGTGACC	7323
9561	Db	AGGACTAGCTGGATTTCTTAGGCTGACTAGAATCCCTAAGCCCTAGCTGGGAAGGTGACC	9620
7324	Qy	ACATCCACCTTTTAAACACGGGGCTTGCAACTTAGYTACACCTGACCCAAATCAGAGAGCTC	7383

Db	9621	ACATCCACCTTTAAACACGGGGCTTGTCACCTTAGCTCACAACCTGACCAATTCAGAGAGCTC	9688
Qy	7384	ACTAAATGCTAAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAATCATYTAATTCGMITG	7443
Db	9681	ACTAAATGCTAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAATCATYTAATTCGMITG	9740
Qy	7444	AGACACAGCAGGAGGAGCAATGATCGGATATAAAACCCAAAGTTCGAGCCCGCAACGG	7503
Db	9741	AGACACAGCAGGAGGAGCAATGATCGGATATAAAACCCAAAGTTCGAGCCCGCAACGG	9800
Qy	7504	CAACCCCTTTGGTCCCTCCCTTTGTATGGAGGCTCTGTTTTCATGCTATTTCACTCT	7563
Db	9801	CAACCCCTTTGGTCCCTCCCTTTGTATGGAGGCTCTGTTTTCATGCTATTTCACTCT	9860
Qy	7564	ATTAATCTTGCARCTGCR	7582
Db	9861	ATTAATCTTGCARCTGCA	9879
RESULT 13			
LOCUS	AR699757		
DEFINITION	Sequence 3 from patent US 6919438.		
ACCESSION	AR699757		
VERSION	AR699757.1	GI:75205546	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 10499)		
AUTHORS	Alliel, P.M., Perin, J.-P. and Rieger, F.		
TITLE	Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses		
JOURNAL	Patent: US 6919438-A 3 19-JUL-2005;		
	Institut National de la Sante et de la Recherche Medical-Inserm;		
	Paris;		
	FR;		
FEATURES			
source	Location/Qualifiers		
	1..10499		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
Query Match	85.9%;	Score 6436;	DB 2; Length 10499;
Best Local Similarity	95.7%;	Pred. No. 0;	
Matches 6680;	Conservative 195;	Mismatches 50;	Indels 54; Gaps 28;
Qy	606	GATGGGAAACGTTCCCGCAAGACAAACAAACGCCCCCTTAAGACGTAATTCGAGAAATTTGGGA	665
Db	2953	GATGGGAAACGTTCCCGCAAGACAAACAAACGCCCCCTTAAGACGTAATTCGAGAAATTTGGGA	3012
Qy	666	MCAATTTGACCTTCAGACACTAGAAGAAACGACTATATTTCTTCGAGTGCCTGCTG	725
Db	3013	CCAATTTGACCTTCAGACACTAGAAGAAACGACTATATTTCTTCGAGTGCCTGCTG	3072
Qy	726	GCACTCTGAGGAAAGTATAAATATAACACCACTTACAGCTAGACATCTTTTGTGTAGAA	785
Db	3073	GCACTCTGAGGAAAGTATAAATATAACACCACTTACAGCTAGACCTCTTTTGTGTAGAA	3132
Qy	786	AAGGCAATGGAGTGAAGTGCCATAAGTACAACTTTCTTTTCAATTAAGAGACAACTCAC	845
Db	3133	AAGGCAATGGAGTGAAGTGCCATAAGTACAACTTTCTTTTCAATTAAGAGACAACTCAC	3192
Qy	846	AATATGTAAAAAGTGTGATTTATGCCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	905
Db	3193	AATATGTAAAAAGTGTGATTTATGCCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	3252
Qy	906	CCAGCATCCCGACTCTTCCCCCAATTAAGGACCCCTTCACCCCAATGGTCCAA	965
Db	3253	CCAGCATCCCGACTCTTCCCCCAATTAAGGACCCCTTCACCCCAATGGTCCAA	3312
Qy	966	AAGGAGATACACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTCCTCCAAATATGA	1025

Db 3313 AAGGAGATAGACAAAGGGTAAACAGTGAAACCAAGAGTGCCAAATATTCCTCCCAATTATGA 3372  
Qy 1026 CCCCTCCCAAGCAGTGGGAGGAGAGAAATTCGGCCCGCAGCAGATGCGATGCTGCTTTTYY 1085  
Db 3373 CCCCT-CCAAGCAGTGGGAGGAGAGAAATTCGGCCCGCAGCAGATGCGATGCTGCTTTTTC 3431  
Qy 1086 TCTCCAGACTTAAAGCAAAATAAAACACAGACTTACGTAAATCTCAGATAAATCCCTGATGG 1145  
Db 3432 TCTCCAGACTTAAAGCAAAATAAAACACAGACTTACGTAAATCTCAGATAAATCCCTGATGG 3491  
Qy 1146 CTATATTTGRTGTTTTACAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGATATATA 1205  
Db 3492 CTATATTTGATGTTTTACAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGATATA-A 3550  
Qy 1206 TGTCTACTGTAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTAATCTGAGCCT 1265  
Db 3551 TGTCTACTGTAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTAATCTGAGCCT 3610  
Qy 1266 GAGRGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGGATANGGATGACAAACAGAAGG 1325  
Db 3611 GAGAGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGG--ATAGGATGACAAACAG-AGG 3667  
Qy 1326 AAGAGAAATGATTCGCCACAGCCAGCAGCAGTTCGCCAGTCTASACCCCTCAATTGGGAC 1385  
Db 3668 AAGAGAAATGATTCGCCACAGCCAGCAGCAGTTCGCCAGTCTAGACCCCTCAATTGG-- 3724  
Qy 1386 ACAGAAATCAGTAAACATGCGGAGATTTGGTCTGCGACACATTTGCTAACTTTGTTGCTASAA 1445  
Db 3725 ACACAGAAATCAGAACAT-GGAGATTTGGTCTGCGACACATTTGCTAACTTTGTTGCTAGAA 3783  
Qy 1446 GGACTAAGGAAACTPAGAGAAATCTAYGAATTTACTCAATGATGTCACCATTAACACA 1505  
Db 3784 GGACTAAGGAAACTPAGAG--AAGTCTATGAATTTACTCAATGATGTCACCATTAACACA 3842  
Qy 1506 CGGGAGGGAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGAGCATTTGAGGAAGC 1565  
Db 3843 -GGGAGGGAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGAGCATTTGAGGAAGC 3901  
Qy 1566 GTGCTCTCTGTGACCTGACTCTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTATCA 1625  
Db 3902 GTGCTCTCTGTGACCTGACTCTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTATCA 3961  
Qy 1626 CTCAGTCAGCTCAGACATTAG-AAAAAATCTCAAAGTCTCGCGTAGCGCCGGAGCAAA 1684  
Db 3962 CTCAGTCAGCTCAGACATTAGAAAAAAATCTCAAAGTCTCGCGTAGCGCCGGAGCAAA 4021  
Qy 1685 ACTTAGAAACCCCTATTGAACTTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC 1744  
Db 4022 ACTTAGAAACCCCTATTGAACTTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC 4081  
Qy 1745 AGGCGGAACAGGACAAACGGGATTAATAAAAGGCCACCGCTTTTATGTCANGACCTCAGG 1804  
Db 4082 AGGCGGAACAGGACAAACGGGATTAATAAAAGGCCACCGCTTTTATGTCANGACCTCAGG 4141  
Qy 1805 CAAAGTGGACTTTGGAGGCTCTGGAAGGGAAGCTGGGCAAAATGATGCTTAATAGG 1864  
Db 4142 CAAAGTGGACTTTGGAGGCTCTGGAAGGGAAGCTGGGCAAAATGATGCTTAATAGG 4201  
Qy 1865 GCTTGCTTCCAGTGGGCTCTACAAGGACACTTTTAAAAAGATTTGCCAAGTAAAGTAAAG 1924  
Db 4202 GCTTGCTTCCAGTGGGCTCTACAAGGACACTTTTAAAAAGATTTGCCAAGTAAAGTAAAG 4261  
Qy 1925 CGGCCCCCTTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 1984  
Db 4262 CGGCCCCCTTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 4321  
Qy 1985 GGACAAAGTCTTTTGAAGTCAGAGCCACTTAACAGATGATCCAGCAGCAGCTAGG 2044  
Db 4322 GGACAAAGTCTCTGAGTCAGAGCCACTTAACAGATGATCCAGCAGCAGCTAGG 4381  
Qy 2045 TGCCTGGGCAAGCGCATCCCATGCTATCCCTCACAAGGCCCTGGGTATGCTTGACC 2104  
Db 4382 TGCCTGGGCAAGCGCATCCCATGCTATCCCTCACAAGGCCCTGGGTATGCTTGACC 4441

Qy 2105 ATTGAGGCCAGGAGGTTGTCCTCTGACACTGCTGGGTCTTCTTAGTCTTACTCTTTC 2164  
Db 4442 ATTGAGGCCAGG-AGGTTGTCCTCTGACACTGCTGGGTCTTCTTAGTCTTACTCTTTC 4500  
Qy 2165 TGTCCCGACAACTGTCTCTCAGATCTCTCACTATTCTGAGGGGTCTCNATAAGAGGGCA 2224  
Db 4501 TGTCCCGACAACTGTCTCTCAGATCTCTCACTA-TCTGAGGGGTCC-TAAGACGGCA 4558  
Qy 2225 GTCACTAGATCTTCTTCTCCAGGCCACTAAGTTATGAATGAACTGGGGAGCTTTATCTTTTCA 2284  
Db 4559 GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTGGGGAGCTTTATCTTTTCA 4616  
Qy 2285 ATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTCTTAGGGAGAGACATTTCTAGCAA 2344  
Db 4617 ATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTCTTAGGGAGAGACATTTCTAGCAA 4676  
Qy 2345 AAGCAGGGCCCAATTATACACCTGAACATAGAGAGAGAAACACCCCTTTGTTGCTTNCCTTC 2404  
Db 4677 AAGCAGGGCCCAATTATACACCTGAACATAGAGAGAGAAACACCCCTTTGTTGCTTNCCTTC 4735  
Qy 2405 CTTGAGGAAGGAAATTAATCTGAACTCTGGGCAACAGAGGACAATATGGAAGGACAAA 2464  
Db 4736 CTTGAGGAAGGAAATTAATCTGAACTCTGGGCAACAGAGGACAATATGGAAGGACAAA 4794  
Qy 2465 GAATGCCCTCTCTGTTCAAGTTTAACTAAAGATTTCCACTTCTTCTTCCCTACCAAGGCA 2524  
Db 4795 GAATGCCCTCTCTGTTCAAGTTTAACTAAAGATTTCCACTTCTTCTTCCCTACCAAGGCA 4854  
Qy 2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTTGTAAGGACTTAAAGC 2584  
Db 4855 GTACCCCTCAGACCCCAAGGCCCAACAGGACTTCCAAAGATTTGTAAGGACTTAAAGC 4914  
Qy 2585 CCAAGGCTTAGTAAACCATGCAATACTCCCTGAGTAATTTCCGTTAGTGGATTTAGGAGG 2644  
Db 4915 CCAAGGCTTAGTAAACCATGCAATACTCCCTGAGTAATTTCCGTTAGTGGATTTAGGAGG 4967  
Qy 2645 CACAGAAACCCAGTGAGGAGGAGTGTAGTGAAGATCTCAGAGATTTATCAATGGAGC 2704  
Db 4968 TACAGAAACCCAGTGAGGAGGAGTGTAGTGAAGATCTCAGAGATTTATCAAT-GAGG 5025  
Qy 2705 GGTGCTCTTTATACCCAGCTGTACTAGCCCTTATATCTGTGTTTCCCAATACCAGA 2764  
Db 5026 TGTGTTCTCTATAGCAGCTGTACTAGCCCTTATATCTGTCTTCCCAATACCAGA 5085  
Qy 2765 GGAAGCAGAGTGGTTTACCTCTGGAACCTTGAAGATGCTTCTTCTGCAATCCCTGTACA 2824  
Db 5086 GGAAGCAGAGTGGTTTACAGTCTTGAAGATGCTTCTTCTGCAATCCCTGTACA 5145  
Qy 2825 TCCTGACTCTCAATTTCTTGTGCTTGAAGATACTTCAACCCAGCATCTCAACTCAG 2884  
Db 5146 TCCTGACTCTCAATTTCTTGTGCTTGAAGATACTTCAACCCAGCATCTCAACTCAG 5205  
Qy 2885 CTGGACTTTTACCCCAAGGTTTCAAGGATAGYCCCATCTATTTGGCCAGGCAATTAG 2944  
Db 5206 CTGGACTTTTACCCCAAGGTTTCAAGGATAGTCCCATCTATTTGGCCAGGCAATTAG 5265  
Qy 2945 CCAAGACTGAGYCAATYMTATCCTCTGACACTTCTTGTCTTCTGTAAGTGGATTT 3004  
Db 5266 CCAAGACTGAGYCAATYMTATCCTCTGACACTTCTTGTCTTCTGTAAGTGGATTT 5323  
Qy 3005 ACTTTTTCGCTGCTTCAAGAACTTGTGCAATCAAGCCAGCCAGCTCTTAAATTT 3064  
Db 5324 ACTTTTTCGCTGCTTCAAGAACTTGTGCAATCAAGCCAGCCAGCTCTTAAATTT 5383  
Qy 3065 CCTCCYACCTGTGCTACAGGTTTCCAAACAPARGCTCACTCTGCTCAGCAGAGT 3124  
Db 5384 CCTCCYACCTGTGCTACAGGTTTCCAAACAPARGCTCACTCTGCTCAGCAGAGT 5443  
Qy 3125 TAAATACCTTAGGCTTAARATTTATCAAGGCAACAGGCCCCCTCAGTGAAGAAATATCA 3184  
Db 5444 ----TACTTAGGCTTAAATTTATCAAGGCAACAGGCCCCCTCAGTGAAGAAATATCA 5499

Qy	3185	GCCTATACTGCGCTTATCCTCATCYCAAAAACCTTAAGCAACTAAGRGRRTTTCCTTGGCRT	3244
Db	5500	GCCTATACTGCGCTTATCCTCATCCAAAAACCTTAAGCAACTAAGGGGATTCCTTGGCGT	5559
Qy	3245	AAVAGGVTTCTGCCGAAWATGGATTCCCGCAGGTGTCBAAATAGCCAGCYCATTAWATA	3304
Db	5560	AATAGGTTTCTGCCGAANAATGGATT-CCCAGGTATGGCAANTAGCCAGGTCAATAANTA	5618
Qy	3305	CASATAATTAAGGAAACTCAGAAAGCCCAATACCCATTTTARTAAAGATGAYAMCTGAAGYMR	3364
Db	5619	CACATAATTAAGGAAACTCAGAAAGCCCAATACCCATTTTAGTAAGATGACAACCTGAAGTAG	5678
Qy	3365	AAGTGGCTTCTCCAGGCCCTTAAGAAAGCCCTTAACCCCAAGYCCCACTGTTAGYTTGCC	3424
Db	5679	AAGTGGCTTCTCCAGGCC------TAACCCAAGCCCACTGTTAGYTTTGGC	5724
Qy	3425	AACRGGGCAAGACTTTTSTTVATAYRTCAAGAAAAAACAAGAAAYAGCTCTRGAGTGCC	3484
Db	5725	AACAGGCAAGACTTTTCTCATATGTTCACAG-AAAAAACAAGNATAGCTCTTAGAGTGCC	5783
Qy	3485	TTACACAGRTCCRAGGAYGAGCTTGCAACCCYRTGGCRYACCTGAAPAGGAAAYTGATG	3544
Db	5784	TTACACAGATCCGAGGAGTACGCTTGCAACCTGTGGCATACCTGACTAAGGAAAATGATG	5843
Qy	3545	TAGTGGCAAGGGTGTGCYCATTTGTTTATVGGGTAGTGGTGGCAGTAGCAGTYKTAGTAT	3604
Db	5844	TAGTGGCAAGGGTGTGACCTCATTTGTTTACGGGTAGTGGTGGCAGTAGCAGCTCTTAGTAT	5903
Qy	3605	CTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGACATCTCATGAKGTGA	3664
Db	5904	CTGAAGCAGTTAAATAATACAGGGAAGAGATCTTACTGTGTGACATCTCATGATGTGA	5963
Qy	3665	AYRCATACTCACTGTAAAAGGAGACTTGTGGCTGTGACAGCAACYGTTTACTTTAAATPTC	3724
Db	5964	ATGCACTACTCACTGTAAAAGGAGACTTGTGGCTGTGACAGCAACYGTTTACTTTAAATGTC	6023
Qy	3725	AGGCTCTATTACTTGAARGGCCAGTGTGTCBACCTGTGCACTTGTGCAACTCTTTAAACCCAG	3784
Db	6024	AGGCTCTATTACTTGAARGGCCAGTGTGTCBACCTGTGCACTTGTGCAACTCTTTAAACCCAG	6083
Qy	3785	YNCATTTCTTCAGACAAATGAGAAAGATAPAAVATRACTGTCAACARTAAATTTCTC	3844
Db	6084	CCACATTTCTTCAGACAAATGAGAAAGATAPAAACATACTGTCAACAGTAAATTTCTC	6143
Qy	3845	AAACCTATGCCACTCGAGGGGACCTTGTAGAGTTTCCTTGTGACTGTATCCYGCCTTCAAC	3904
Db	6144	AAACCTATGCCACTCGAGGGGACCTTGTAGAGTTTCCTTGTGACTGTATCCCGACC-TCAC	6202
Qy	3905	TTGTATACTGATGGAAGTTCTTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGCAGTG	3964
Db	6203	TTGTATACTGATGGAAGTTCTTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGCAGTG	6262
Qy	3965	GTCAGTGATAAATGGATATTTGAAGTAAATCCCTCACTCCAGGAAGTAGTGTAGCTTR	4024
Db	6263	GTCAGTGATAAATGGATATTTGAAGTAAATCCCTCACTCCAGGAAGTAGTGTAGCTTR	6322
Qy	4025	GCAGAACTAATAGCCYTCATTKGGGCACTAGAAATTAGGACAAGRAAAAAGGGYAAATATA	4084
Db	6323	GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTAGGACAAGRAAAAAGGGCAATATA	6382
Qy	4085	TATACAGACTCTRTARTATGTACCTTAGTCNTCCATGCCCATCMACATCAGGAAGCCATTAGBARATTA	4204
Db	6383	TATACAGACTCTAAATATGTCTTACCTAGTCCTCATGCCCATGCAGCAATATGGAAGAA	6442
Qy	4145	AGGGAATTCCTAATCTCYGARGNAACCTTATCMACATCAGGAAGCCATTAGBARATTA	4204
Db	6443	AGGGAATTCCTAATCTCTGAGAGAACCTTATCAACATCAGGAAGCCATTAGGAANTTA	6502
Qy	4205	TTATVGGCWGTACAGAAACCTTARAGAGGTGGMAGTCTTTACACTGCTGGGTGCATCANAA	4264
Db	6503	TTATVGGCWGTACAGAAACCTTAAAGAGGTGGCAGTCTTTACACTGCGCGGGTGCATCAGAA	6562
Qy	4265	GGAAAGPAAAGGGAAATASAGRGAAYTGCACAGCAKATATTGAAGTAAAGAGAGCTGCA	4324

Db 7643 CATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCAITTTTATCTACCCCAACTGCGG 7702  
Qy 5404 TTAAGTGGCTGGAGTGAAGTCTTGGATACATCACACTTGAGTCAATCTCTGGTACTGCG 5463  
Db 7703 TTAAGTGGCTGGAGTGAAGTCTTGGATACATCACACTTGAGTCAATCTCTGGTACTGCG 7762  
Qy 5464 CAAGGAACCTGAAATCTCAGGAGCAACGCTAGTCTATTCTCTGTGAACCTCTAGAGGATT 5523  
Db 7763 CAAGGAACCTGAAATCTCAGGAGCAACGCTAGTCTATTCTCTGTGAACCTCTAGAGGATT 7822  
Qy 5524 TCGCGCTCTCTTCAACAACAACAGGAGGAAGTAACTAAATCATAAATCTCCCCATG 5583  
Db 7823 TGGCGCTCTCTTCAACAACAACAGGAGGAAGTAACTAAATCATAAAT-CCCCATG 7881  
Qy 5584 GSCCTCCCTATCATATTTTCTCTKTAATGTTTSTTTTACCTSTTTTCACTCTCACTGCA 5643  
Db 7882 GSCCTCCCTATCATATTTTCTCTKTAATGTTTSTTTTACCTSTTTTCACTCTCACTGCA 7941  
Qy 5644 CCCCTCCCATGCGCTGTATGACCAAGTAGTCCCTCTYACMAGAGTTCCTATGGAGAATG 5703  
Db 7942 CCCCTCCCATGCGCTGTATGACCAAGTAGTCCCTCTTACCAAGAGTTCCTATGGAGAATG 8001  
Qy 5704 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCACC 5763  
Db 8002 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACC 8061  
Qy 5764 TTCACTGCGCCACACCCATATGCCCCCGGAACTGTCTATCACTCTGCGCACTCTTTGCGATGAT 5823  
Db 8062 TTCACTGCGCCACACCCATATGCCCCCGGAACTGTCTATCACTCTGCGCACTCTTTGCGATGAT 8121  
Qy 5824 GCAATATCTCATATTGGACAGGAATAATGATTAATCTCTAGTTGCTCTGGAGGACTTGA 5883  
Db 8122 GCAATATCTCATATTGGACAGGAATAATGATTAATCTCTAGTTGCTCTGGAGGACTTGA 8181  
Qy 5884 GTCACTGTCTGTGGACTTACTTTTCAACCCAACTGTGTATGTCTGTATGCTGGGGGTGAGTTCAA 5943  
Db 8182 GTCACTGTCTGTGGACTTACTTTTCAACCCAACTGTGTATGTCTGTATGCTGGGGGTGAGTTCAA 8241  
Qy 5944 GATCAGGCAAGAGAAAAACATGTAAAGAAATCTCTCCAACTCAACCCGGGTACATGCG 6003  
Db 8242 GATCAGGCAAGAGAAAAACATGTAAAGAAATCTCTCCAACTCAACCCGGGTACATGCG 8301  
Qy 6004 ACCTTAGCCCCACAGGAGTACATCTCTCAAACTACATGAACCTCTCGTACCCAT 6063  
Db 8302 ACCTTAGCCCCACAGGAGTACATCTCTCAAACTACATGAACCTCTCGTACCCAT 8361  
Qy 6064 ACTCGCCTGGTAAGCTATTTAATACCACTCTCACTGGGCTCCATGAGGTCTCGGCCCAA 6123  
Db 8362 ACTCGCCTGGTAAGCTATTTAATACCACTCTCACTGGGCTCCATGAGGTCTCGGCCCAA 8421  
Qy 6124 AACCCTACTAACTGTGGATATGCTCCCTCGAACTTCAGGCCATATGTTTCAATTCCT 6183  
Db 8422 AACCCTACTAACTGTGGATATGCTCCCTCGAACTTCAGGCCATATGTTTCAATTCCT 8481  
Qy 6184 GTACCTGAACATGGAACAACTTCAGCAGAAATAAACACCACTTCGGTTTTTACTAGGA 6243  
Db 8482 GTACCTGAACATGGAACAACTTCAGCAGAAATAAACACCACTTCGGTTTTTACTAGGA 8541  
Qy 6244 CCTCTGTGTTTCCAAATGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
Db 8542 CCTCTGTGTTTCCAAATGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 8601  
Qy 6304 AATACTACATACACAACTCCCAATGCGATCAGGTGGGTAACTCTCTCCCAACAATA 6363  
Db 8602 AATACTACATACACAACTCCCAATGCGATCAGGTGGGTAACTCTCTCCCAACAATA 8661  
Qy 6364 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATGCTGTTTGTGAT 6423  
Db 8662 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATGCTGTTTGTGAT 8721  
Qy 6424 GGTCTTTCAAGATCTATGTGCTTCTCTCAATCTTAGTGCCGCCCATGRCCTACTACACT 6483  
Db 8722 GGTCTTTCAAGATCTATGTGCTTCTCTCAATCTTAGTGCCGCCCATGRCCTACTACACT 8781

Qy 6484 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCT 6543  
Db 8782 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCT 8841  
Qy 6544 CCTTTTGTATAGGAGCAGGAGTCTAGGTGCACCTAGGTACTTGGCAITTTGGCGGTATACA 6603  
Db 8842 CCTTTTGTATAGGAGCAGGAGTCTAGGTGCACCTAGGTACTTGGCAITTTGGCGGTATACA 8901  
Qy 6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGAATGGAACGGGTC 6663  
Db 8902 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGAATGGAACGGGTC 8961  
Qy 6664 GCCGACTCCCTGGTCACTCTGCAAGATCAACTTAACCTCTAGCAGCAGTACTCTTCBA 6723  
Db 8962 GCCGACTCCCTGGTCACTCTGCAAGATCAACTTAACCTCTAGCAGCAGTACTCTTCBA 9021  
Qy 6724 AATCGAAGAGCTTTAGACTYGTCTAACCGCTGARAGAGGGGAACCTGTTTATTTTAGGG 6783  
Db 9022 AATCGAAGAGCTTTAGACTYGTCTAACCGCTGARAGAGGGGAACCTGTTTATTTTAGGG 9081  
Qy 6784 GAAGAATGCTGTTTATTTATTTAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 9082 GAAGAATGCTGTTTATTTAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 9141  
Qy 6844 CSAGATCGAATACAACTGAKAGCAGAGGCTTCGAAACACTTGGACCTGGGGCCTCTCT 6903  
Db 9142 CGAGATCGAATACAACTGAGAGCAGAGGCTTCGAAACACTTGGACCTGGGGCCTCTCT 9201  
Qy 6904 AGCRAATGGATGCTCTGGATTCTCCCTCTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
Db 9202 AGCRAATGGATGCTCTGGATTCTCCCTCTCTTAGGACCTCTAGCAGCTATATATTTGCTA 9261  
Qy 6964 CTCTCTTTTGGACCTCTGATCTTTTAACTCTTGTGTAACTTTTCTCTTCAGAAATCGAA 7023  
Db 9262 CTCTCTTTTGGACCTCTGATCTTTTAACTCTTGTGTAACTTTTCTCTTCAGAAATCGAA 9321  
Qy 7024 GCTGTAAACTACAAATGGAGCCCAAGATGAGTCCCAAGACTTAAGTCTACCGCAGACCC 7083  
Db 9322 GCTGTAAACTACAAATGGAGCCCAAGATGAGTCCCAAGACTTAAGTCTACCGCAGACCC 9381  
Qy 7084 CTGACCCGCTGTAGCCCAAGATGAGTCTTAACTGATCATCAAAAGGCAACCTCTCTGAG 7143  
Db 9382 CTGACCCGCTGTAGCCCAAGATGAGTCTTAACTGATCATCAAAAGGCAACCTCTCTGAG 9441  
Qy 7144 GAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGGAAGCAAGTTAGAGCGGT 7203  
Db 9442 GAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGGAAGCAAGTTAGAGCGGT 9501  
Qy 7204 SGTGGCCCAACCTCCCAACAGCACTTAGTCTTCTCTGAGATGGGGGACTGAGAGAC 7263  
Db 9502 C-TGGGCCCAACCTCCCAACAGCACTTAGTCTTCTCTGAGATGGGGGACTGAGAGAC 9560  
Qy 7264 AGGACTAGCTGGATTCTCTAGGCTGATTAAGAAATCCYTAAGCTTAGTGGGAAGGTGACC 7323  
Db 9561 AGGACTAGCTGGATTCTCTAGGCTGATTAAGAAATCCYTAAGCTTAGTGGGAAGGTGACC 9620  
Qy 7324 ACATCCAACCTTTAAACACGGGCTTGGCAACTTAGVTCAACCTGACCAATCAGAGGCTC 7383  
Db 9621 ACATCCAACCTTTAAACACGGGCTTGGCAACTTAGVTCAACCTGACCAATCAGAGGCTC 9680  
Qy 7384 ACTAAATGCTAAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCCAATCATYTTATTCMTG 7443  
Db 9681 ACTAAATGCTAAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCCAATCATYTTATTCMTG 9740  
Qy 7444 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAGTCTTCGAGCCGGCAACCG 7503  
Db 9741 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAGTCTTCGAGCCGGCAACCG 9800  
Qy 7504 CAACCCCTTTGGGTCCCTCTCTTTGATGAGGAGCTCTGTTTTTCAATGCTATTTCTCTCT 7563  
Db 9801 CAACCCCTTTGGGTCCCTCTCTTTGATGAGGAGCTCTGTTTTTCAATGCTATTTCTCTCT 9860

QY 7564 ATTAATCTTGCACCTGCR 7582  
 Db |||||||:|||||:  
 9861 ATTAATCTTGCACCTGCA 9879

RESULT 14  
 AX007980 10499 bp DNA linear PAT 06-SEP-2000  
 DEFINITION Sequence 3 from Patent WO9967395.  
 ACCESSION AX007980  
 VERSION AX007980.1 GI:9995677  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Perin, J.P., Rieger, F. and Alliel, P.M.  
 TITLE Nucleic sequence and deduced protein sequence family with human  
 endogenous retroviral motifs, and their uses  
 JOURNAL Patent: WO 9967395-A 3 29-DEC-1999;  
 INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER  
 FRANCOIS (FR); ALLIEL PATRICK M (FR)  
 FEATURES  
 source 1..10499  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 85.9%; Score 6436; DB 2; Length 10499;  
 Best Local Similarity 95.7%; Pred. No. 0;  
 Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

QY 606 GATGGGAAACGTTCCCGCAAGACAAACGCGCCCTAAGACGTAATCTGCGARAAATGGGA 665  
 Db |||||||:|||||:  
 2953 GATGGGAAACGTTCCCGCAAGACAAACGCGCCCTAAGACGTAATCTGCGARAAATGGGA 3012

QY 666 MCAATTGACCCCTCAGACACTAAGAAAGAACGACTTATATCTCTGCGAGTCCCGCCTG 725  
 Db :|||||:  
 3013 CCAATTGACCCCTCAGACACTAAGAAAGAACGACTTATATCTCTGCGAGTCCCGCCTG 3072

QY 726 GCACTCTGAGGAAAGTAAATATATAACACCATCTTACAGTAGACATCTTTTGTAGAA 785  
 Db GCACCTCTGAGGAAAGTAAATATATAACACCATCTTACAGTAGACATCTTTTGTAGAA 3132

QY 786 AAGGCAATGGAGTGAAGTCCCATAGTACAAATCTTCTTCAATTAAGAGACAATCTCAC 845  
 Db |||||||:|||||:  
 3133 AAGGCAATGGAGTGAAGTCCCATAGTACAAATCTTCTTCAATTAAGAGACAATCTCAC 3192

QY 846 AATTATGTAAGAAAGTGTATTTATGCGCTACAGGAAGCCCTTACAGAGTCTACCTCCCTATC 905  
 Db |||||||:|||||:  
 3193 AATTATGTAAGAAAGTGTATTTATGCGCTACAGGAAGCCCTTACAGAGTCTACCTCCCTATC 3252

QY 906 CCAGCATCCCGACTCTCTCCCAATTAATAAGGACCCCGCTTCAACCCAAATGGTCCAA 965  
 Db |||||||:|||||:  
 3253 CCAGCATCCCGACTCTCTCCCAATTAATAAGGACCCCGCTTCAACCCAAATGGTCCAA 3312

QY 966 AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAATATGA 1025  
 Db AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAATATGA 3372

QY 1026 CCCTCCCAAGCAGTGGAGGAAGAAATTCGGCCAGCAGAGTGCATGTGCTTTT 1085  
 Db |||||||:|||||:  
 3373 CCCTCCCAAGCAGTGGAGGAAGAAATTCGGCCAGCAGAGTGCATGTGCTTTT 3431

QY 1086 TCTCCAGACTTAAGCAATAAAGACAGACTTTAGGTAAATCTCAGATAAATCTGATGG 1145  
 Db |||||||:|||||:  
 3432 TCTCCAGACTTAAGCAATAAAGACAGACTTTAGGTAAATCTCAGATAAATCTGATGG 3491

QY 1146 CTATATGTTTACAGGGTTAGGACAAATCTTTGATCTGACATGGAGATATATA 1205  
 Db |||||||:|||||:  
 3492 CTATATGTTTACAGGGTTAGGACAAATCTTTGATCTGACATGGAGATATATA 3550

Db 3492 CTATATGTTTACAGGGTTAGGACAAATCTTTGATCTGACATGGAGATATA-A 3550  
 QY 1206 TGTCACTGCTAAATCAGACACATAACCCAAATCAGAGAACTGCACCATAACTGCAGCCT 1265  
 Db |||||||:|||||:  
 3551 TGTCACTGCTAAATCAGACACATAACCCAAATCAGAGAACTGCACCATAACTGCAGCCT 3610

QY 1266 GAGGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGGATANGATGACACAGAAGG 1325  
 Db |||||||:|||||:  
 3611 GAGAGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGG-ATAGGATGACACAG-AGG 3667

QY 1326 AAAGANAATGATTCCCCACAGGCCAGCAGCTCCAGTCTTASACCCCTCATTTGGGAC 1385  
 Db |||||||:|||||:  
 3668 AAAGAGAAATGATTCCCCACAGGCCAGCAGCTTCCAGTCTAGACCCCTCATTTGG- 3724

QY 1386 ACAGAAATCAGTAAATGGAGATTGGTCTGTCAGACATTTGCTAACTTTGCTGTASAA 1445  
 Db |||||||:|||||:  
 3725 ACACAGAAATCAGAAACAT-GGAGATTGGTCTGTCAGACATTTTGTAACTTTGCTGTAGAA 3783

QY 1446 GGACTAAGGAAACTAGGAAGAAATCTAYGAATTAATCAATGATCTCCACCATTAACACA 1505  
 Db |||||||:|||||:  
 3784 GGACTAAGGAAACTAGGAAG-AAGTCTATGAATTAATCAATGATGTCCACCATTAACACA 3842

QY 1506 GGGGAAAGGAAGAAATCTTACTGCTCTTCTGGAGAGACTTAAGGGAGGCAATTCAGGAAGC 1565  
 Db |||||||:|||||:  
 3843 -GGGAAAGGAAGAAATCTTACTGCTCTTCTGGAGAGACTTAAGGGAGGCAATTCAGGAAGC 3901

QY 1566 GTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 1625  
 Db |||||||:|||||:  
 3902 GTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 3961

QY 1626 CTCAGTCAGTCGACACATTAG-AAAGAACTTCAAAAGTCTGCGTAGGCCCCGGAGGAAA 1684  
 Db |||||||:|||||:  
 3962 CTCAGTCAGTCGACACATTAGAAAGAACTTCAAAAGTCTGCGTAGGCCCCGGAGGAAA 4021

QY 1685 ACTTAGAAACCTTATTTGAACCTTGGCAACTTGGCACTTGGTATTAATAGAGATCAGAGGAGC 1744  
 Db |||||||:|||||:  
 4022 ACTTAGAAACCTTATTTGAACCTTGGCACTTGGTATTAATAGAGATCAGAGGAGC 4081

QY 1745 AGCGGAAACAGGACAAACGGGATTTAAAAAGGCGCACCGCTTTAGTCTATGACCTCAGG 1804  
 Db |||||||:|||||:  
 4082 AGCGGAAACAGGACAAACGGGATTTAAAAAGGCGCACCGCTTTAGTCTATGACCTCAGG 4141

QY 1805 CAAATGGAATTTGGAGGCTCTGGAAAGGAAAGCTGGGCAAAATGAAATGCTTAATAGG 1864  
 Db |||||||:|||||:  
 4142 CAAATGGAATTTGGAGGCTCTGGAAAGGAAAGCTGGGCAAAATGAAATGCTTAATAGG 4201

QY 1865 GCTTGTCTCCAGTGGCTCTACAGGACACTTTTAAAAAGATTTGTCAGTAGAGTAAGTAAG 1924  
 Db |||||||:|||||:  
 4202 GCTTGTCTCCAGTGGCTCTACAGGACACTTTTAAAAAGATTTGTCAGTAGAGTAAGTAAG 4261

QY 1925 CGCGCCCTTCTGTCATGCGCTTTATTTCAAGGGAATCACTTGGAGGCGCCCACTGCCCCAGG 1984  
 Db |||||||:|||||:  
 4262 CGCGCCCTTCTGTCATGCGCTTTATTTCAAGGGAATCACTTGGAGGCGCCCACTGCCCCAGG 4321

QY 1985 GGACAAAGGTCTTTTGAAGTCAAGGACCTTAACAGATGATCCAGCAGCAGGACTGAGGG 2044  
 Db |||||||:|||||:  
 4322 GGACAAAGGTCTCTTGAAGTCAAGGACCTTAACAGATGATCCAGCAGCAGGACTGAGGG 4381

QY 2045 TGCCTGGGCAAGCGGCATCCATGCGCATCACTCCAGAGCCCTGGGTATGCTTTGACC 2104  
 Db |||||||:|||||:  
 4382 TGCCTGGGCAAGCGGCATCCATGCGCATCACTCCAGAGCCCTGGGTATGCTTTGACC 4441

QY 2105 ATTGAGGGGCAAGAGGTTGCTCTTGGACACTTGGTGGCGCTTCTTACTTCTTACTCTTC 2164  
 Db |||||||:|||||:  
 4442 ATTGAGGGGCAAG-AGGTTGCTCTTGGACACTTGGTGGCGCTTCTTACTTCTTCTTC 4500

QY 2165 TGTCCCGGCAAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGACGGGCA 2224  
 Db |||||||:|||||:  
 4501 TGTCCCGGCAAACTGTCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA 4558

QY 2225 GTCACTAGATCTTTTCCAGCCACTAAGATTATGAATCTGGGAGGCTTTATTTCTTTTTCAC 2284  
 Db |||||||:|||||:  
 4559 GTCACTAGATCTTTTCCAGCCACTAAGATTATG-ACTGGGGAGGCTTTATTTCTTTTTCAC 4616

QY 2285 ATGCTTTTCTAATATGCTTGAAGCCCACTACCTGTTAGGGAGACATCTTAGCAA 2344  
DB 4617 ATGCTTTTCTAATATGCTTGAAGCCCACTACCTGTTAGGGAGACATCTTAGCAA 4676  
QY 2345 AAGCAGGGCCCAATTATACACCTTGAACATAGAGAAAGAACACCCCTGTTGTTNCCCTG 2404  
DB 4677 AAGCAGGGCCCAATTATACACCTTGAACATAGAGAAAGAACACCCCTGTTGTTGTT-CCCTG 4735  
QY 2405 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAAAGCAATATGGACGACCAAA 2464  
DB 4736 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAAAGCAATATGGACGAC-CAAA 4794  
QY 2465 GAATCCCGTCTCTGTTCAAGTTAACTAAAGATTCCACTTCCCTTCCCTACCAAGGCA 2524  
DB 4795 GAATCCCGTCTCTGTTCAAGTTAACTAAAGATTCCACTTCCCTTCCCTACCAAGGCA 4854  
QY 2525 GTACCCCTCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTGTTAAGGACTTAAAAAGC 2584  
DB 4855 GTACCCCTCTCAGACCCCAAGGCCCAACAGGACTTCCAAAGATTGTTAAGGACTTAAAAAGC 4914  
QY 2585 CCAAGCCTTAGTAAACCAATCAATCTCCCTGCAAGTAAATCCGTTAGTGGATTTAGGAGG 2644  
DB 4915 CCAAGCCTTAGTAAACCAATCAATCTCCCTGCAAGTAAATCCGTTAGTGGATTTAGGAGG 4967  
QY 2645 CACAGAAACCCAGTGGACAGTGGAGGTTAGTGGAGATCTCAGGATTTCAATGGAGGC 2704  
DB 4968 TACAGAAACCCCAACAGCAGTGGAG-GGTTAGTGGAGATCTCAGGATTTCAAT-AGGCG 5025  
QY 2705 CGTTGCTCTTTATACCAACGCTGACTAGCCCTTATATCTGTMVTTCCCAATAACAGAG 2764  
DB 5026 TGTGTTCTCTATAGCCAGCTGATACCTTAGCCCTTATCTCTGCTTTCCCAATAACAGAG 5085  
QY 2765 GGAAGCAGAGTGGTTTACASTCCTCGACCTTGAAGATGCTTCTGATCCCTGTATACA 2824  
DB 5086 GGAAGCAGAGTGGTTTACAGTCTCGACCTTCAGGATGCTTCTGATCCCTGTATACA 5145  
QY 2825 TCCTGACTCTCAATTTCTGTTTGGCTTTGAAGATACCTTCAACCCCAACATCTCAACTCAC 2884  
DB 5146 TCCTGACTCTCAATTTCTGTTTGGCTTTGAAGATACCTTCAACCCCAACATCTCAACTCAC 5205  
QY 2885 CTGGACTTCTTTACCCCAAGGCTTCAGGATAGYCCCATCTATTTTGGCCAGGCATTAGC 2944  
DB 5206 CTGGACTTCTTTACCCCAAGGCTTCAGGATAGTCCCATCTATTTTGGCCAGGCATTAGC 5265  
QY 2945 CCAAGACTTGAGCCCAATCCTCATACCTGGACA-CTTGTCTCTCGGTAGGTGGATGATTT 3004  
DB 5266 CCAAGACTTGAGCCCAATCCTCATACCTGGACA-CTTGTCTCTCGGTAGGTGGATGATTT 5323  
QY 3005 ACTTTTGGYGCYRTTCAGAAACCTTTGGCCATCAGCCACCCCAAGCCTCTTAAATTT 3064  
DB 5324 ACTTTTGGCCGCCATTCAGAAACCTTTGGCCATCAGCCACCCCAAGCCTCTTAAATTT 5383  
QY 3065 CTTCCGYACTCTGTGCTACAWGGTTTCCAAAGCABARGCTCACTCTGCTCAGCAGGT 3124  
DB 5384 CTTCCGCTACTCTGTGGCTACATGGTTTCCAAACCAAGGCTCACTCTGCTCAGCAGGT 5443  
QY 3125 TAAATACTTAGGCTTAARATTTACAAAGGCCACCGGCCCTTCAGTGAGGAAYRYATCCA 3184  
DB 5444 ---TACTTAGGGCTAAATTTATCCAAAGGCCACCGGCCCTTCAGTGAGGAACACATCCA 5499  
QY 3185 GCCTATACTGGCTTATCTCTCATYCAAAACCCCTAAAGCACTAAGRBRTTCTTGGCCT 3244  
DB 5500 GCCTATACTGGCTTATCTCTCATYCAAAACCCCTAAAGCACTAAGRBRTTCTTGGCCT 5559  
QY 3245 AAYAGYTTCTCCGAAAWATGATTTCCCGAGTGTGGCRAAATAGCCAGGYCATTAWATA 3304  
DB 5560 AATAGGTTCTCCGAAAWATGATTT-CCCGAGTGTGGCRAAATAGCCAGGTCAITTAATA 5618  
QY 3305 CASTAATTAAGGAATCTCAGAAAGCCAATACCACTTARTATAGATGGAYAMCTGAAGYMR 3364  
DB 5619 CACTAATTAAGGAATCTCAGAAAGCCAATACCACTTARTATAGATGGACAACTGAAGTAG 5678

QY 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCCAAGTGTAAAGYTTGCC 3424  
DB 5679 AAGTGGCTTTCCAGGCCCTTAAACCCCAAGYCCCAAGTGTAAAGYTTGCC 5724  
QY 3425 AACRGGGCAAGACTTTTATYATRTACAGAAAAAACAAGAAAYAGCTCTTGGAGTCC 3484  
DB 5725 AACAGGGCAAGACTTTTCTTCATATGTCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCC 5783  
QY 3485 TTACACAGRTCCRAGGGAYGAGCTTGCNACCYRTGGCBRYACTGASTAAGGAAATGTATG 3544  
DB 5784 TTACACAGATCCGAGGGATGAGCTTGCNACCTTGGCATACCTGCTAAGGAAATTTGATG 5843  
QY 3545 TAGTGGCAAGGGTTGRCYTCATTTGTTAYGGGTAGTGGCTGCTAGCAGTYKTAGTAT 3604  
DB 5844 TAGTGGCAAGGGTTGACCTCATTTGTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT 5903  
QY 3605 CTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
DB 5904 CTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGATGTGA 5963  
QY 3665 AYRGCATACTCAGCTTAAAGGAGACTTGTGGCTGTGACAGAAACVGTGTTACTTAATRTC 3724  
DB 5964 ATGCGCATACTCAGCTTAAAGGAGACTTGTGGCTGTGACAGAAACVGTGTTACTTAATRTC 6023  
QY 3725 AGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAACCCAG 3784  
DB 6024 AGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAACCCAG 6083  
QY 3785 YCNATTTCTTCCAGACAATGAAGAAAGATPARAAYATACTGTGCAACAARTAAATTTCTC 3844  
DB 6084 CCACATTTCTTCCAGACAATGAAGAAAGATPARAAYATACTGTGCAACAARTAAATTTCTC 6143  
QY 3845 AAACTATGCACTCGAGGGGACCTTTTAGARGTTCCVTTGACTGATCCYGACCTTCAAC 3904  
DB 6144 AAACTATGCACTCGAGGGGACCTTTTAGARGTTCCVTTGACTGATCCYGACCTTCAAC 6202  
QY 3905 TTGTATATCTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATCAGTG 3964  
DB 6203 TTGTATATCTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATCAGTG 6262  
QY 3965 GTCAGTGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACTAGTGTGCTYAGCTR 4024  
DB 6263 GTCAGTGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACTAGTGTGCTYAGCTR 6322  
QY 4025 GCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTTAGGAGAAAGRAAAAGGYAATATA 4084  
DB 6323 GCAGAACTAATAGCCCTCACTTGGGCACCTAGAAATTTAGGAGAAAGRAAAAGGYAATATA 6382  
QY 4085 TATCAGACTCTARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSRAGAA 4144  
DB 6383 TATCAGACTCTARTATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSRAGAA 6442  
QY 4145 AGGGAATTTCTTAACCTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204  
DB 6443 AGGGAATTTCTTAACCTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTAGGARATTA 6502  
QY 4205 TTAATGCGWGTACAGAAACCTTARAGAGGTGMAGTCTTACACTGCGYGGGTCTCANAAA 4264  
DB 6503 TTAATGCGWGTACAGAAACCTTARAGAGGTGMAGTCTTACACTGCGYGGGTCTCANAAA 6562  
QY 4265 GGAAAGRAAGGGGAAATAAAGRAAAATGCGCAAGCAKATATTGAAGCMAAAGAGCTGCA 4324  
DB 6563 GGAAAGRAAGGGGAAATAAAGRAAGAACTGCGCAAGCAKATATTGAAGCMAAAGAGCTGCA 6622  
QY 4325 AGGCAGGACCTCTCAATTTAGAAATGCTTATTAACCTTCCCTTTAGTATAGGGTAAATCCCTTC 4384  
DB 6623 AGGCAGGACCTCTCAATTTAGAAATGCTTATTAACCTTCCCTTTAGTATAGGGTAAATCCCTTC 6682  
QY 4385 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGG-CAG 4443  
DB 6683 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGGAGAG 6742  
QY 4444 TTTTCTCCCTCGGGAACCGTTAGCCACTGAAGAGGGGAAATACTTTTGGCTTGCACACTAT 4503



|||||  
6743 TTTTCTCCCTCGGACGGCTAGCCACTGAAGAGGGAATAATCTTTTGCTCGAACTAT 6802  
QY  
4504 CCAATGGAATTAATTAAACCCCTTCATCAAACTTTCACTTAGGCATCGATAGCACCCA 4563  
Db  
6803 CCAATGGAATTAATTAAACCCCTTCATCAAACTTTCACTTAGGCATCGATAGCACCCA 6862  
QY  
4564 TCARATGGCCAAATCAATTTACTGACACAGCCCTTTTCAAACTATCAAGCARATAKT 4623  
Db  
6863 TCAGATGGCCAAATCAATTTACTGACACAGCCCTTTTCAAACTATCAAGCAGATAGT 6922  
QY  
4624 CAGGGCTGTGAAKTGTGCCARAAATAATCCCTGCGCTTATCGCCAAAGCTCCTTCAGG 4683  
Db  
6923 CAGGGCTGTGAAKTGTGCCAGAAATAATCCCTGCGCTTATCGCCAAAGCTCCTTCAGG 6982  
QY  
4684 ARAACAAARAACAGGCCATTACCTGRRARAACCTGGCAACTGATTTTACCCACAAGCCC 4743  
Db  
6983 AGAACAAARAACAGGCCATTACCTGRRARAACCTGGCAACTGATTTTACCCACAAGCCC 7042  
QY  
4744 AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATATTTTCAGGGTTGGGCARAG 4803  
Db  
7043 AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATATTTTCAGGGTTGGGCARAG 7102  
QY  
4804 GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGCTAATAAAGGCACTAGTTTCATGAATA 4863  
Db  
7103 GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGCTAATAAAGGCACTAGTTTCATGAATA 7162  
QY  
4864 ATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGTTTCCAGGCC 4923  
Db  
7163 ATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGTTTCCAGGCC 7222  
QY  
4924 ACAGTAACCCAGGAGTATCCAGGGCTTAGGTATAGTATACATTAACCTGCGCTGA 4983  
Db  
7223 ACAGTAACCCAGGAGTATCCAGGGCTTAGGTATAGTATACATTAACCTGCGCTGA 7282  
QY  
4984 AGGCCACAGTCTCCGAGGAGTGCAGAAAATGAATGAAYACTCAAGGACATCTAAA 5043  
Db  
7283 AGGCCACAGTCTCCGAGGAGTGCAGAAAATGAATGAAYACTCAAGGACATCTAAA 7342  
QY  
5044 AAGCAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGCTATAGCCCTTAAAAAGA 5103  
Db  
7343 AAGCAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGCTATAGCCCTTAAAAAGA 7402  
QY  
5104 ATCTGCAATTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGGCC 5163  
Db  
7403 ATCTGCAATTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGGCC 7462  
QY  
5164 TTCAATAACCAATGACCTTGCTTGACCCAGACGCAACTTAGTTGAGACATCACCT 5223  
Db  
7463 TTCAATAACCAATGACCTTGCTTGACCCAGACGCAACTTAGTTGAGACATCACCT 7522  
QY  
5224 CCTTAGCCAAATATCAACAGTTCTTAAACATTACAAGGAACCTATCCTCTGAGAAGGG 5283  
Db  
7523 CCTTAGCCAAATATCAACAGTTCTTAAACATTACAAGGAACCTATCCTCTGAGAAGGG 7582  
QY  
5284 GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCCTCTAAATCCC 5343  
Db  
7583 GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCCTCTAAATCCC 7642  
QY  
5344 CATCCCTAGATCATCTGGGAGGACCTTACCAGTCAATTTATTTACCCCACTGGCG 5403  
Db  
7643 CATCCCTAGATCATCTGGGAGGACCTTACCAGTCAATTTATTTATCTACCCCACTGGCG 7702  
QY  
5404 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAAGTCAATCTGATCTGC 5463  
Db  
7703 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAAGTCAATCTGATCTGC 7762  
QY  
5464 CAAAGAACTGAAATCCAGGAGACAACGCTAGCTATTCTGTGAACTCTAGAGGATT 5523  
Db  
7763 CAAAGAACTGAAATCCAGGAGACAACGCTAGCTATTCTGTGAACTCTAGAGGATT 7822  
QY  
5524 TGGCCCTGCTTCAAAACAAACAGGAGGAAGTAATCAATCAATAATCCCCCATG 5583  
|||||

Db  
7823 TGGCCCTGCTCTTCAAAACAAACACAGGAGGAAGTAATCAATAATCATATAAT-CCCCATG 7881  
QY  
5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTAYACCTSTTTTCACTCTCACTGCA 5643  
Db  
7882 GCCTCCCTTATCATATTTTCTCTTACTGTTTCTTTTACCCTCTTCACTCTCACTGCA 7941  
QY  
5644 CCCCCTCAATGCGCTGTATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGGAGATG 5703  
Db  
7942 CCCCCTCAATGCGCTGTATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGGAGATG 8001  
QY  
5704 CAGCGTCCCGGAAATATTCATGCGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAACC 5763  
Db  
8002 CAGCGTCCCGGAAATATTCATGCGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAACC 8061  
QY  
5764 TTCACTGCCCACACCCATATGCGCCCAACTGCTATCACTCTGCGCACTCTTTGCACTGCAT 5823  
Db  
8062 TTCACTGCCCACACCCATATGCGCCCAACTGCTATCACTCTGCGCACTCTTTGCACTGCAT 8121  
QY  
5824 GCNAATACTCATTTATGGACAGGAAATAATTAATCTAGTTGCTCTGGAGACTTGGGA 5883  
Db  
8122 GCNAATACTCATTTATGGACAGGAAATAATTAATCTAGTTGCTCTGGAGACTTGGGA 8181  
QY  
5884 GTCACCTGCTGTTGGACTTTTACCCCAAACTGGTATGCTGATGGGGTGGAGTTCAA 5943  
Db  
8182 GTCACCTGCTGTTGGACTTTTACCCCAAACTGGTATGCTGATGGGGTGGAGTTCAA 8241  
QY  
5944 GATCAGGCAAGGAAAAATGTAAAGAAAGTAATCTCCCAACTCAACCGGGTACATGGC 6003  
Db  
8242 GATCAGGCAAGGAAAAATGTAAAGAAAGTAATCTCCCAACTCAACCGGGTACATGGC 8301  
QY  
6004 ACTCTAGCCCCCTAACAGGACTAGATCTCTCAAACTATAGTAAACCTCTCCGTACCCAT 6063  
Db  
8302 ACTCTAGCCCCCTAACAGGACTAGATCTCTCAAACTATAGTAAACCTCTCCGTACCCAT 8361  
QY  
6064 ACTCGCTGCTGAAGCTATTTAATACCCCTCACTGGGCTCCATGAGTCTCGGCCCCA 6123  
Db  
8362 ACTCGCTGCTGAAGCTATTTAATACCCCTCACTGGGCTCCATGAGTCTCGGCCCCA 8421  
QY  
6124 AACCTTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGGCCATATGTTTCAATCCCT 6183  
Db  
8422 AACCTTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGGCCATATGTTTCAATCCCT 8481  
QY  
6184 GTACTGAAACATGGAACAACTTTCAGCACAGAAATAAACAACCTTCCTGTTTGTAGGGA 6243  
Db  
8482 GTACTGAAACATGGAACAACTTTCAGCACAGAAATAAACAACCTTCCTGTTTGTAGGGA 8541  
QY  
6244 CCTCTGTTTCCAAATSTGGAATACCCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
Db  
8542 CCTCTGTTTCCAAATSTGGAATACCCATACCTCAAACTCACCTGTGTAAATTTAGC 8601  
QY  
6304 AATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAATA 6363  
Db  
8602 AATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAATA 8661  
QY  
6364 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCCTATGCTGTTTGTGAAT 6423  
Db  
8662 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCCTATGCTGTTTGTGAAT 8721  
QY  
6424 GSCTCTTCAGATCTATGTGCTTCTCTCATTTAGTGGCCCCCATGTCATCTACACT 6483  
Db  
8722 GSCTCTTCAGATCTATGTGCTTCTCTCATTTAGTGGCCCCCATGTCATCTACACT 8781  
QY  
6484 GAAACAGATTTTATACAGTTATGTCTATCTAGAGCCCGCAACAAAGAGTACCCATCTT 6543  
Db  
8782 GAAACAGATTTTATACAGTTATGTCTATCTAGAGCCCGCAACAAAGAGTACCCATCTT 8841  
QY  
6544 CTTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTACTTGGCAATTCGGCGGTATCA 6603  
Db  
8842 CTTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTACTTGGCAATTCGGCGGTATCA 8901  
QY  
6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 6663  
Db  
8902 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 8961





[illegible]

Db	4137	GGGCTTGCTTCAGTGCAGTCTCAAGAGCACATTGGAAGAGTGTGCAAGTAGAAGTA	4193
Qy	1923	AGCGCCCCCTGCTGCATGCCCCCTTATTTCAAGAGGAATCACTCGAAGGCCACCTGCGCCCA	1982
Db	4197	AGCGCCCCCTGCTGCATGCCCCCTTATTTCAAGAGGAATCACTCGAAGGCCACCTGCGCCCA	4256
Qy	1983	GGGACAAGAGTCTTTTGTAGTCAGAAGCCACTAACACAGATGATCCAGCAGCAGGACTGAG	2042
Db	4257	GGGACAAGAGTCTCTGTAGTCAGAAGCCACTAACACAGATGATCCAGCAGCAGGACTGAG	4316
Qy	2043	GGTGCTGGGGCAGCGCCATCCCATGCCATCACTTCACAGAGCCCTGGGTATGTTGA	2102
Db	4317	GGTGCTGGGGCAGCGCCATCCCATGCCATCACTTCACAGAGCCCTGGGTATGTTGA	4376
Qy	2103	CCATTGAGGGCCAGGAAGTGTGCTCCTGAGCACTGGTGGGTCTTCTTAGTCTTACTCT	2162
Db	4377	CCATTGAGGGCCAGG-AGGTTGTCTCTGAGCACTGGTGGGTCTTCTTAGTCTTACTCT	4435
Qy	2163	TCTGTCCGGCAAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGG	2222
Db	4436	TCTGTCCGGCAAACTGTCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACAG	4493
Qy	2223	CAGTCACTAGTACTTTTCCGACGCCATGAAGTATGAACCTGGGGAGCTTTATCTTTTC	2282
Db	4494	CAGTCACTAGTACT-:TTCTCCGACGCCATAAGTTATG-ACTGGGGAGCTTTATCTTTTC	4551
Qy	2283	ACATGCTTTTCTAATTATGCTGAAAGCCCCACTACTCTGTTAGGGAGAGACATCTAGC	2342
Db	4552	ACATGCTTTTCTAATTATGCTGAAAGCCCCACTACTCTGTTAGGGAGAGACATCTAGC	4611
Qy	2343	AAAAGCAGGGGCCATTATACACCTGCACTAGGAGAAGGAACAACCCGTTTGTGTGCCCC	2402
Db	4612	AAAAGCAGGGGCCATTATACACCTGCACTAGGAGAAGGAACAACCCGTTTGTGT-CCCC	4670
Qy	2403	TGCTTGAGGAAGGAATTAACTCTGAAGTCTGGGCAACAGAGAACAATATGACGAGCCA	2462
Db	4671	TGCTTGAGGAAGGAATTAACTCTGAAGTCTGGGCAACAGAGAACAATATGACGAG-CA	4729
Qy	2463	AAGAATGCCCGTCTCTGTTCAAGTTAACTAAAGGATTCACATTCCTTCCCTACCAAGG	2522
Db	4730	AAGAATGCCCATCTCTGTTCAAGTTAACTAAAGGATTCACATTCCTTCCCTACCAAGG	4789
Qy	2523	CAGTACCCCTCTCAGACCCCAAGGCCAACAGGATTCCAAAGATTTGTAAGGACTTAAAA	2582
Db	4790	CAGTACCCCTCTCAGACCCCAAGGCCAACAGGACTTCAAAGATTTGTAAGGACCTAAAA	4849
Qy	2583	GCCCAAGGCTTAGTAAACCATGCATAACTCCCTGCAGTAAATTCGTAGTGGATTGAGA	2642
Db	4850	GCCCAAGGCTTAGTAAAACTATGCAGTAAACCCCTGCAGTACTCCAAT-----TTTAGGAG	4904
Qy	2643	GGCACAAACCAGTGGAGCAGTGGAGGGTTAGTGCAAGATCTCAGGATTTATCAATGGAG	2702
Db	4905	ATTACAAACCACACAGACAGTGGGA-GGTTAGTGCAAGATCTCAGGATTTATCAAT-GAG	4962
Qy	2703	GCGCTTCTCTTTTATACCCAGCTGTACCTAGCCCTTATACTGTGMVTTTCCCAATATCCA	2762
Db	4963	GCTGTTGTTCTCTATAGCCAGCTGTACCTAGCCCTTATACTGTGCTTTCCCAATATCCA	5022
Qy	2763	GAGAAACAGAGTGGTTTACASTCTGGACCTTMAGGATGCCCTTTCTTGCATCCCTGTA	2822
Db	5023	GAGAAACAGAGTGGTTTACAGTCTCTGGACCTTTCAGGATGCGCTTCTTCTGCATCCCTGTA	5082
Qy	2823	CATCCTGACTCTCAATCTTGTTTGGCTTTGAAGTACTTCAAAACCCATCTCAATC	2882
Db	5083	CATCCTGACTCTCAATCTTGTTTGGCTTTGAAGTACTTCAAAACCCATCTCAATC	5142
Qy	2883	ACCTGGACTTTTATCCCAAGGGTTTCAGGATAGYCCCACTCTATTTTGGCCAGCATTTA	2942
Db	5143	ACCTGGACTTTTATCCCAAGGGTTTCAGGATAGTCCCACTCTATTTTGGCCAGCATTTA	5202
Qy	2943	GCCCAAGACTTGAGYCARVTMTCATACCTGGACACTCTTTGTCTCTTCRGTAAGTGGATGAT	3002
Db	5203	GCCCAAGACTTGAGCAATCTCATACCTGGACACTCTTGCTCTCTTCGGTAGTGGATGAT	5262



Db	7402	 CCTTCATAACCAATGACCTTGTGCTTGA - CCAAGACAGCCAACTTAGTTGCAGACATCAC
Qy	5222	CTCCTTAGCCAAATATCAACAAGTCTTTAAAAACATTACAAGAAACCTTATCCCTGAGAGA
Db	7461	CTCCTTAGCCAAATATCAACAAGTCTTTAAACATTACAAGAAACCTTATCCCTGAGAGA
Qy	5282	GGGAAAAGAACTATTCCACCCWGTGTGACATGGTATTAGTCAAGTCCCTTCVCTTAATTC
Db	7521	GGGAAAAGAACTATTCCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC
Qy	5342	CCCATCCCTAGATACATCTCTGGGAAGAGACCTACCCAGTCAATTTATTTATACCCCACTGC
Db	7581	CCCATCCCTAGATACATCTCTGGGAAGAGACCTACCCAGTCAATTTATTTATACCCCACTGC
Qy	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCCTGGATACT
Db	7641	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTCGAGTCAAAACCTTGGATACT
Qy	5462	GCCAAAGGAACCTGAAAATCCAGAGACAACGCTAGCTATTCTCTGTGAAACCTCTAGAGGA
Db	7701	GCCAAAGGAACCTGAAAATCCAGAGACAACGCTAGCTATTCTCTGTGAAACCTCTAGAGGA
Qy	5522	TTTTCGCGCTGCTCTTTCAACAACAACAGAGAGGAAGTAACATAAAATCATAAATCCCCCA
Db	7761	TTTTCGCGCTGCTCTTTCAACAACAACAGAGAGGAAGTAACATAAAATCATAAAT - ACCCA
Qy	5582	TGGSCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTVAACCTSTTTTCACCTCTCACTG
Db	7820	TGGCCCTCCCTTATCATATTTTCTCTTTACTGTTCTTTTACCTCTTTTCACCTCTCACTG
Qy	5642	CACCCCTCCATGCGCGTGTATGACCACTAGTCTCCCTTACCAAGAGTTTCTATGAGAA
Db	7880	CACCCCTCCATGCGCGTGTATGACCACTAGTCTCCCTTACCAAGAGTTTCTATGAGAA
Qy	5702	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCCA
Db	7940	TGCAGCGTCCCGGAAATATTGATGCCCATCATATAGGAGTCTTTCTTAAGGGAACCCCCA
Qy	5762	CTTTCACTGCCACACCCCATATGCCCGCAACTGCCTATCACTCTGCCACCTCTTTTGATGC
Db	8000	CTTTCACTGCCACACCCCATATGCCCGCAACTGCCTATCACTCTGCCACCTCTTTTGATGC
Qy	5822	ATGCAATACTCATATTATGGACAGGAAAAATGATTAACTCTAGTGTCTCTGGAGGACTTG
Db	8060	ATGCAATACTCATATTATGGACAGGAAAAATGATTAACTCTAGTGTCTCTGGAGGACTTG
Qy	5882	GAGTCACTGCTGTGTGACTTACTTCAACCAACTGGTATGCTCTGATGGGGGTGAGTTCC
Db	8120	GAGTCACTGCTGTGTGACTTACTTCAACCAACTGGTATGCTCTGATGGGGGTGAGTTCC
Qy	5942	AAGATCAGGCAAGAGAAAAACATGTAAAAAGATTAATCTCCCAACTCACCCGGGTACATG
Db	8180	AAGATCAGGCAAGAGAAAAACATGTAAAAAGATTAATCTCCCAACTCACCCGGGTACATG
Qy	6002	GCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAAGAACCTCCGTAACC
Db	8240	GCGCCTCTAGGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAAGAACCTCCGTAACC
Qy	6062	ATACTCGCCTGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC
Db	8300	ATACTCGCCTGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC
Qy	6122	AAAAACCTACTAATCTGTGTGATATGCTCCCTCCCTGAACCTTCARGCCCATATGTTTCAATCC
Db	8360	AAAAACCTACTAATCTGTGTGATATGCTCCCTCCCTGAACCTTCARGCCCATATGTTTCAATCC
Qy	6182	CTGTACTGAACAAATGGAAACAACTTCAGCAAGAAAAATAAACACCACTTCGGTTTTAGTAG
Db	8420	CTGTACTGAACAAATGGAAACAACTTCAGCAAGAAAAATAAACACCACTTCGGTTTTAGTAG
Qy	6242	GACCTCTTGTTTCCAATSTGGAATAACCATACCTCAAACTCAACTGTTGTAATAATTTA

DB	8480	GACCTCTTGTTTCCAAATCTGGAATAACCCATACTACCTCAAACTCAGCTGCTGTAATAATTTA	8533
QY	6302	GCAATACTACATACACAACCAACTCCCAATGSCATCAGTGGTAACCTCCTCCACACAAA	6361
DB	8540	GCAATACTACAGACAACCAACTCCCAATGSCATCAGTGGTAACCTCCTCCACACAAA	8599
QY	6362	TAGTCTGCTACCTCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTGA	6421
DB	8600	TAGTCTGCTCCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTGA	8659
QY	6422	ATGGCTCTTCAGATCTATGTGCTTCTCTCATCTTGTAGTCCGCCCYATGRCCATCTACA	6481
DB	8660	ATGGCTCTTCAGATCTATGTGCTTCTCTCATCTTGTAGTCCGCCCTATGACCATCTACA	8719
QY	6482	CTGAACAAGATTTATACAGTTATGTCTATCTAAGCCGCCCAACAAAGAGTACCCATTC	6541
DB	8720	CTGAACAAGATTTATACAATATATGTCTATCTAAGCCGCCCAACAAAGAGTACCCATTC	8779
QY	6542	TTCCTTTTGTATAGGACGAGTGTCTAGGTGCACTAGTACTGTGCAATTTGGCGGTATCA	6601
DB	8780	TTCCTTTTGTATAGGACGAGTGTCTAGGTGCACTAGTACTGTGCAATTTGGCGGTATCA	8839
QY	6602	CAACCTCTACTCAGTTCTACTACAACACTATCTCAAGACTAATAATGGGACATGGAAACGG	6661
DB	8840	CAACCTCTACTCAGTTCTACTACAACACTATCTCAAGAACTAATAATGGGACATGGAAACGG	8899
QY	6662	TCGCCGACTCCCTGGTCACTCTCAAGACTCAACTTAACTCCCTAGCAGCAGTAGTCTCTTC	6721
DB	8900	TCGCCGACTCCCTGGTCACTCTGCAAGATCACTTAACTCCCTAGCAGCAGTAGTCTCTTC	8959
QY	6722	RAAATCGAAGAGCTTTTAGACTYGTCAACCGCTGARAGAGGGGAACTGTGTTATTTTTAG	6781
DB	8960	RAAATCGAAGAGCTTTTAGACTYGTCAACCGCTGAAGAGGGGAACTGTGTTATTTTTAG	9019
QY	6782	GGGAAGAATGCTGTTATATGTTAACTCCGGATTCGTCACTGAGAAAGTTTAAAGAAA	6841
DB	9020	GGGAAGAATGCTGTTATATGTTAACTCCGGATTCGTCACTGAGAAAGTTTAAAGAAA	9079
QY	6842	TTCSAGATCGAATAACAACGTAKAGCAGAGAGCTTCGAACACTCGAACCTCGGGGCTCC	6901
DB	9080	TTCSAGATCGAATAACAACGTAGAGAGAGAGCTTCGAACACTCGAACCTCGGGGCTCC	9139
QY	6902	TCAGCCRAATGGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATGC	6961
DB	9140	TCAGCCRAATGGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATGC	9199
QY	6962	TACTCTCTTTGGACCTGTATCTTTTAACTCCTCTGTTTAACTTTTGTCTCTTCAGAAATCG	7021
DB	9200	TACTCTCTTTGGACCTGTATCTTTTAACTCCTCTGTTTAACTTTTGTCTCTTCAGAAATCG	9259
QY	7022	AAGCTGTAAACTACAAATCGAGCCCAAGATGCAGTCCAAAGACTAAGACTTACCGCAGAC	7081
DB	9260	AAGCTGTAAACTACAAATCGAGCCCAAGATGCAGTCCAAAGACTAAGACTTACCGCAGAC	9319
QY	7082	CCCTGGACCGGCTGYTAGCCCAAGTCTGATGTTTAAATGACATCAAAGGACACCCCTCCTG	7141
DB	9320	CCCTGGACCGGCTGYTAGCCCAAGTCTGATGTTTAAATGACATCAAAGGACACCCCTCCTG	9379
QY	7142	AGGAAATCTCAGCTGCACAACCTCTACTACGCCCCCAATTCAGCAGGAAGAGTTAGAGCG	7201
DB	9380	AGGAAATCTTAACTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGAGTTAGAGCG	9439
QY	7202	GTGTGCGGCCAACTCCCAACAGCACTTAGGTTTTCTCTGTTGAGATGGGGACCTAGAG	7261
DB	9440	GTGTGCGGCCAACTCCCAACAGCACTTAGGTTTTCTCTGTTGAGATGGGGACCTAGAG	9499
QY	7262	ACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAATCCYTAAGCCTTAGTGTGGGAAGGTGA	7321
DB	9500	ACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAATCCYTAAGCCTTAGTGTGGGAAGGTGA	9559
QY	7322	CCACATCCACTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGC	7381
DB	9560	CCACATCCACTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAGC	9619

Qy 7382 TCACCTAAATGCTAATTAGGCAAGACAGGAGGTAAAGAAATAGCCAAATCATYTTATTGCM 7441  
 Db 9620 TCACCTAAATGCTAATTAGGCAAGAAACAGGAGGTAAAGAAATAGCCAAATCATCTATTGCC 9679  
 Qy 7442 TGAGAGCACAGCAGGAGGACAAATGATCGGGATATAAACCCCAAGTTCGAGCCGGCAAC 7501  
 Db 9680 TGAGAGCACAGCAGGAGGACAAATGATCGGGATATAAACCCCAAGTTCGAGCTGGCAAC 9739  
 Qy 7502 GGCAACCCCTTTGGGTCCCTTCTTGTATGGAGCTCTGTTTCATGCTATTTCAC 7561  
 Db 9740 GGCAACCCCTTTGGGTCCCTTCTTGTATGGAGCTCTGTTTCATGCTATTTCAC 9799  
 Qy 7562 CTATTAAATCTTGCAACTGCR 7582  
 Db 9800 CTATTAAATCTTGCAACTGCA 9820

Search completed: August 2, 2006, 03:16:35  
 Job time : 27475 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 15:08:51 ; Search time 5532 Seconds  
(without alignments)  
16841.079 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatcggtataaacc.....tattaatcttgactgtgr 7582

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq.\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq.\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7489.6	100.0	7582	8	US-10-632-793-30
2	7489.6	100.0	7582	9	US-10-717-580-11
3	6448	86.1	56093	3	US-09-873-367C-81
4	6448	86.1	56093	10	US-10-843-641A-81
5	6436	85.9	10499	13	US-11-028-539-3
6	6124.6	81.8	8523	3	US-09-854-867-21
7	6124.6	81.8	8523	10	US-10-786-970A-21
8	3738.4	49.9	6394	10	US-10-450-763-20014
9	3431	45.8	4349	10	US-10-450-763-12279
10	3309.6	44.2	7466	10	US-10-450-763-12279
11	3105.4	41.5	177175	15	US-11-121-086-79
12	3105.4	41.5	189993	15	US-11-121-086-78
13	3049.8	40.7	22436	3	US-09-997-722-148
14	2948	39.4	139573	12	US-10-506-513-3
15	2948	39.4	156416	10	US-10-461-862-12
16	2932	39.1	1514	10	US-10-450-763-3413
17	2889.6	38.6	2930	3	US-09-902-535-1

18	2888.6	38.6	2946	6	US-10-114-893-134	Sequence 134, Appl
19	2888.6	38.6	2946	6	US-10-016-249-3	Sequence 3, Appl
20	2734.4	36.5	3831	10	US-10-450-763-7531	Sequence 7531, Appl
21	2688.6	35.9	2938	8	US-10-632-793-21	Sequence 21, Appl
22	2688.6	35.9	2938	9	US-10-717-580-2	Sequence 2, Appl
23	2613.2	34.9	3372	8	US-10-632-793-28	Sequence 28, Appl
24	2613.2	34.9	3372	9	US-10-717-580-9	Sequence 9, Appl
25	2550.4	34.1	2599	13	US-11-028-539-1	Sequence 1, Appl
26	2541.6	33.9	4535	10	US-10-450-763-12009	Sequence 12009, A
27	2368.4	31.6	8279	10	US-10-450-763-12278	Sequence 12278, A
28	2282.4	30.5	2784	13	US-11-028-539-4	Sequence 4, Appl
29	2251.8	30.1	2372	8	US-10-632-793-29	Sequence 29, Appl
30	2251.8	30.1	2372	9	US-10-717-580-10	Sequence 10, Appl
31	2241	29.9	8294	10	US-10-450-763-20013	Sequence 20013, A
32	2220.4	29.6	161334	6	US-10-087-192-730	Sequence 730, Appl
33	2216	29.6	46340	13	US-11-028-539-66	Sequence 66, Appl
34	2204.6	29.4	2782	8	US-10-632-793-26	Sequence 26, Appl
35	2204.6	29.4	2782	9	US-10-717-580-7	Sequence 7, Appl
36	2201.4	29.4	2782	8	US-10-133-036-1	Sequence 1, Appl
37	2201.4	29.4	3633	10	US-10-821-234-543	Sequence 543, Appl
38	2168	28.9	8399	3	US-09-854-867-26	Sequence 26, Appl
39	2168	28.9	8399	10	US-10-786-970A-26	Sequence 26, Appl
40	2131.4	28.5	2575	9	US-10-717-580-14	Sequence 14, Appl
41	2049.4	27.4	3205	10	US-10-450-763-8040	Sequence 8040, Appl
42	2032	27.1	2055	13	US-11-028-539-22	Sequence 22, Appl
43	2026.6	27.1	2585	9	US-10-717-580-13	Sequence 13, Appl
44	1982.2	26.5	3867	10	US-10-450-763-4425	Sequence 4425, Appl
45	1982.2	26.5	3867	10	US-10-450-763-12268	Sequence 12268, A

## ALIGNMENTS

### RESULT 1

```

US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARAMHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1309)..(1309)

```

OTHER INFORMATION: n = a or g or c or t/u		
FEATURE:		
NAME/KEY: misc feature		
LOCATION: (1331)..(1331)		
OTHER INFORMATION: n = a or g or c or t/u		
FEATURE:		
NAME/KEY: misc feature		
LOCATION: (2213)..(2213)		
OTHER INFORMATION: n = a or g or c or t/u		
FEATURE:		
NAME/KEY: misc feature		
LOCATION: (2398)..(2398)		
OTHER INFORMATION: n = a or g or c or t/u		
FEATURE:		
NAME/KEY: misc feature		
LOCATION: (3787)..(3787)		
OTHER INFORMATION: n = a or g or c or t/u		
FEATURE:		
NAME/KEY: misc feature		
LOCATION: (4115)..(4115)		
OTHER INFORMATION: n = a or g or c or t/u		
FEATURE:		
NAME/KEY: misc feature		
LOCATION: (4261)..(4261)		
OTHER INFORMATION: n = a or g or c or t/u		
US-10-632-793-30		
Query Match 100.0%; Score 7489.6; DB 8; Length 7582;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	CAACAAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGAGCCCCCTTTGGGTCC 60
DB	1	CAACAAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGAGCCCCCTTTGGGTCC 60
QY	61	CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTTCACTTAAATCTTCAACTGCA 120
DB	61	CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTTCACTTAAATCTTCAACTGCA 120
QY	121	CTCTTCGTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180
DB	121	CTCTTCGTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180
QY	181	TGTTTGCCACCAACCGCANACTGCGCTGACTCCCATCCCTCTGGATCTCGAGGCTGC 240
DB	181	TGTTTGCCACCAACCGCANACTGCGCTGACTCCCATCCCTCTGGATCTCGAGGCTGC 240
QY	241	CGCTGTGCTCTGATCCAGGARGCGCCATTCGCGCTCCCAATTTGGGCTTAAGGCTTG 300
DB	241	CGCTGTGCTCTGATCCAGGARGCGCCATTCGCGCTCCCAATTTGGGCTTAAGGCTTG 300
QY	301	CATTGTNCTGCAACGGCTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTANTCACT 360
DB	301	CATTGTNCTGCAACGGCTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTANTCACT 360
QY	361	GGGTTCCATGGTTCTCTTCTGTGACCCACCGCTTCTTAATAKAACTATAACACTTACCA 420
DB	361	GGGTTCCATGGTTCTCTTCTGTGACCCACCGCTTCTTAATAKAACTATAACACTTACCA 420
QY	421	TGGCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAAGCACTCCAGGTCAGAGAAATAC 480
DB	421	TGGCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAAGCACTCCAGGTCAGAGAAATAC 480
QY	481	GARGCTTGGCACCACATCTTGGAAAGCGCTGCTACCTCTTGGAAAGTGGTTTCCACCACATC 540
DB	481	GARGCTTGGCACCACATCTTGGAAAGCGCTGCTACCTCTTGGAAAGTGGTTTCCACCACATC 540
QY	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGTRACATTTTGGGTRACCGGACATCC 600
DB	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGTRACATTTTGGGTRACCGGACATCC 600
QY	601	MAAGTCAATGGAAACCTTCCCGCAAGACAAAACCCCTTAGACGTATTTCTGGARAAT 660
DB	601	MAAGTCAATGGAAACCTTCCCGCAAGACAAAACCCCTTAGACGTATTTCTGGARAAT 660
QY	661	TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTGCAGTGCC 720
DB	661	TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTGCAGTGCC 720
QY	721	GCCTGGCACTCTCGAGGAACTATATAATTATAACCACTCTTACAGCTAGACATCTTTTG 780
DB	721	GCCTGGCACTCTCGAGGAACTATATAATTATAACCACTCTTACAGCTAGACATCTTTTG 780
QY	781	TAGAAAAGGCAAAATGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 840
DB	781	TAGAAAAGGCAAAATGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 840
QY	841	CTCAAAATTTATATAAAAGTGTATTTATGCTTACAGAAAGCTTTCAGAGTCTACCTCC 900
DB	841	CTCAAAATTTATATAAAAGTGTATTTATGCTTACAGAAAGCTTTCAGAGTCTACCTCC 900
QY	901	CTATCCAGCATCCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCCAATGG 960
DB	901	CTATCCAGCATCCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCCAATGG 960
QY	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTATTTCCCAAT 1020
DB	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTATTTCCCAAT 1020
QY	1021	TATGACCCCTCCCAAGCAGTGGAGGAGAAATTTGGCCCAAGCAGAGTGCATGTGCT 1080
DB	1021	TATGACCCCTCCCAAGCAGTGGAGGAGAAATTTGGCCCAAGCAGAGTGCATGTGCT 1080
QY	1081	TTTTTCTCCAGACTTAAAGCAATTAAGCAACTTGGTAAATTTCTCAGATAAATCT 1140
DB	1081	TTTTTCTCCAGACTTAAAGCAATTAAGCAACTTGGTAAATTTCTCAGATAAATCT 1140
QY	1141	GATGCTATATTTGTTTCAAGGGTTAGGCAATTTCTTTGATCTGACATGGAGAGAT 1200
DB	1141	GATGCTATATTTGTTTCAAGGGTTAGGCAATTTCTTTGATCTGACATGGAGAGAT 1200
QY	1201	ATATATGTCATGCTTAAATCAGACACTAAACCCCAATGAGAAAGTGCACCACTGC 1260
DB	1201	ATATATGTCATGCTTAAATCAGACACTAAACCCCAATGAGAAAGTGCACCACTGC 1260
QY	1261	AGCTGAGGTTTGGCGATCTCTGATCTCAGTCACTGATGATGATGATGATGATGATGATGAT 1320
DB	1261	AGCTGAGGTTTGGCGATCTCTGATCTCAGTCACTGATGATGATGATGATGATGATGATGAT 1320
QY	1321	GAAAGAAAGAAATGATTTCCCAAGGCGAGCAGCTTCCAGTCTTASACCTCATTTG 1380
DB	1321	GAAAGAAAGAAATGATTTCCCAAGGCGAGCAGCTTCCAGTCTTASACCTCATTTG 1380
QY	1381	GGGACACAGAAATCAGTAACTGGAGATTTGGTGTGAGACATTTGCTTAACTGTGTGC 1440
DB	1381	GGGACACAGAAATCAGTAACTGGAGATTTGGTGTGAGACATTTGCTTAACTGTGTGC 1440
QY	1441	TASAGGACTAAGGAAACTAGGAAATCTTAAGAAATCTTAAGAAATCTTAAGAAATCTTAAG 1500
DB	1441	TASAGGACTAAGGAAACTAGGAAATCTTAAGAAATCTTAAGAAATCTTAAGAAATCTTAAG 1500
QY	1501	ACACAGGCGAAGGAAAGAAATCTTACTGCTTCTGAGAGACTTAAGGAGGCAATTTGAG 1560
DB	1501	ACACAGGCGAAGGAAAGAAATCTTACTGCTTCTGAGAGACTTAAGGAGGCAATTTGAG 1560
QY	1561	GAAAGCTGCTCTCTGTCACTGACTCTTCTGAAAGCCAACTAACTTTAAAGCGTAAAGTT 1620
DB	1561	GAAAGCTGCTCTCTGTCACTGACTCTTCTGAAAGCCAACTAACTTTAAAGCGTAAAGTT 1620
QY	1621	TATCACTCAGTCACTGAGCACTTAAAGAAACTTCAAAAGCTTCCCTAGCCCGGAG 1680
DB	1621	TATCACTCAGTCACTGAGCACTTAAAGAAACTTCAAAAGCTTCCCTAGCCCGGAG 1680
QY	1681	CAAAACTTAGAAACCTTATTTGCACTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACT 1740
DB	1681	CAAAACTTAGAAACCTTATTTGCACTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACT 1740



QY 1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGAGGCAACCGCTTTAGTCATGACCCCT 1800  
DB 1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGAGGCAACCGCTTTAGTCATGACCCCT 1800  
QY 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGGAAGAAAGCTGGGCAAAATGAATGCTTAA 1860  
DB 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGGAAGAAAGCTGGGCAAAATGAATGCTTAA 1860  
QY 1861 TAGGGCTTGCTCCAGTGGGCTTACAGGACACTTTTAAAAAGATGTCCAAAGTAGAAG 1920  
DB 1861 TAGGGCTTGCTCCAGTGGGCTTACAGGACACTTTTAAAAAGATGTCCAAAGTAGAAG 1920  
QY 1921 TAAGCCGCCCTTCGTCCTCAAGCCCTTATTTCAAGGGAATCACTGGAAAGGCCCACTGCC 1980  
DB 1921 TAAGCCGCCCTTCGTCCTCAAGCCCTTATTTCAAGGGAATCACTGGAAAGGCCCACTGCC 1980  
QY 1981 CAGGGGACAAAGGTCTTTTGAAGTCAGAAAGCCACTAACAGATGATCCAGCAGCAGACTG 2040  
DB 1981 CAGGGGACAAAGGTCTTTTGAAGTCAGAAAGCCACTAACAGATGATCCAGCAGCAGACTG 2040  
QY 2041 AGGGTGCTGGGCAAGCGCATCCCATGCGCATCACTTCCAGAGCCCTTGGGTATGCTT 2100  
DB 2041 AGGGTGCTGGGCAAGCGCATCCCATGCGCATCACTTCCAGAGCCCTTGGGTATGCTT 2100  
QY 2101 GACCATTTAGGGCCAGGAAGGTGTCTCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
DB 2101 GACCATTTAGGGCCAGGAAGGTGTCTCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
QY 2161 CTTCTCTCCGGACAACTGTCTCCAGATCTGTCACTATTCTCAGGGGGTCCNTAAGAGC 2220  
DB 2161 CTTCTCTCCGGACAACTGTCTCCAGATCTGTCACTATTCTCAGGGGGTCCNTAAGAGC 2220  
QY 2221 GGCACTCAGTACTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTT 2280  
DB 2221 GGCACTCAGTACTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTT 2280  
QY 2281 TCACATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTGTAGGAGAGACATTTCTA 2340  
DB 2281 TCACATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTGTAGGAGAGACATTTCTA 2340  
QY 2341 GCAAAAGCAGGGGCCAATTATACACCTGAACATAGGAGAGGACACCCGTTTGTGTGTC 2400  
DB 2341 GCAAAAGCAGGGGCCAATTATACACCTGAACATAGGAGAGGACACCCGTTTGTGTGTC 2400  
QY 2401 CCTGCTTGAGGAGGAATTAATCCTGAACTCTGGGCAACAGAGGACAAATATGGACGAGC 2460  
DB 2401 CCTGCTTGAGGAGGAATTAATCCTGAACTCTGGGCAACAGAGGACAAATATGGACGAGC 2460  
QY 2461 CAAAGAAATCCCGTCTCTGTTCAAGTTAAACTAAAGGATTCACCTCTCTTTCCCTACCAA 2520  
DB 2461 CAAAGAAATCCCGTCTCTGTTCAAGTTAAACTAAAGGATTCACCTCTCTTTCCCTACCAA 2520  
QY 2521 GGCAGTACCCCTCAGACCCAGGCCCCACAGGATTCCAAAGATTTAGGACTTAA 2580  
DB 2521 GGCAGTACCCCTCAGACCCAGGCCCCACAGGATTCCAAAGATTTAGGACTTAA 2580  
QY 2581 AAGCCCAAGGCTTAGTAAACCAATGCAATACTCCCTGCAAGTAACTCCGTAGTGGATTGAG 2640  
DB 2581 AAGCCCAAGGCTTAGTAAACCAATGCAATACTCCCTGCAAGTAACTCCGTAGTGGATTGAG 2640  
QY 2641 GAGGCAACAGAAACCCAGTGGAGAGGTTAGTGCAAGATTCAGGATTTATCAATGG 2700  
DB 2641 GAGGCAACAGAAACCCAGTGGAGAGGTTAGTGCAAGATTCAGGATTTATCAATGG 2700  
QY 2701 AGGCCGTTCTCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGTTTCCCAATAC 2760  
DB 2701 AGGCCGTTCTCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGTTTCCCAATAC 2760  
QY 2761 CAGAGGAACAGAGTGGTTTACASTCTCGACCTTMAGGATGCTTCTCTCATCCCTG 2820  
DB 2761 CAGAGGAACAGAGTGGTTTACASTCTCGACCTTMAGGATGCTTCTCTCATCCCTG 2820

QY 2821 TACATCCTGACTCTCAATTTCTTTGTTGGCTTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
DB 2821 TACATCCTGACTCTCAATTTCTTTGTTGGCTTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
QY 2881 TCACCTTGGACTTTTTTACCCCAAGGCTTCAGGAGATAGYCCCCATCTATTTGGCCAGCAT 2940  
DB 2881 TCACCTTGGACTTTTTTACCCCAAGGCTTCAGGAGATAGYCCCCATCTATTTGGCCAGCAT 2940  
QY 2941 TAGCCCAAGACTTGAAGYCACTYMTCACTCTGGACACTCTTGTCTTCTTCTGTAGTGGATG 3000  
DB 2941 TAGCCCAAGACTTGAAGYCACTYMTCACTCTGGACACTCTTGTCTTCTTCTGTAGTGGATG 3000  
QY 3001 ATTTACTTTTTRGCGCCYFTTCAGAAACCTTGTGCCATCAAGCCACCCAGCBCTCTTMA 3060  
DB 3001 ATTTACTTTTTRGCGCCYFTTCAGAAACCTTGTGCCATCAAGCCACCCAGCBCTCTTMA 3060  
QY 3061 ATTTCTCTCGYACCTGTGGCTACAWGTTTCCAAACSAARGCTCACTCTGTCTCACAGC 3120  
DB 3061 ATTTCTCTCGYACCTGTGGCTACAWGTTTCCAAACSAARGCTCACTCTGTCTCACAGC 3120  
QY 3121 AGGTTAAATACTTAGGCTTAARATTTATCCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRA 3180  
DB 3121 AGGTTAAATACTTAGGCTTAARATTTATCCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRA 3180  
QY 3181 TCCAGCCTATACCTGGCTTATCTCATCYCAAAACCCCTAAAGCMACTAAGRBRRTTCTTGG 3240  
DB 3181 TCCAGCCTATACCTGGCTTATCTCATCYCAAAACCCCTAAAGCMACTAAGRBRRTTCTTGG 3240  
QY 3241 GRTTAAYAGGTTCTGCGGAATGATTTCCCGAGTTCAGGTTGGGCAATAGCCAGGYCATTA 3300  
DB 3241 GRTTAAYAGGTTCTGCGGAATGATTTCCCGAGTTCAGGTTGGGCAATAGCCAGGYCATTA 3300  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTARTTAAGATGGAYAMCTGAA 3360  
DB 3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTARTTAAGATGGAYAMCTGAA 3360  
QY 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCTTAAACCCCAAGYCCCAAGTGTTAAGYT 3420  
DB 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCTTAAACCCCAAGYCCCAAGTGTTAAGYT 3420  
QY 3421 TGCCACACGGGCAAGACTTTTSTTATATVATVATVATVATVATVATVATVATVATVATVAT 3480  
DB 3421 TGCCACACGGGCAAGACTTTTSTTATATVATVATVATVATVATVATVATVATVATVATVAT 3480  
QY 3481 GTCTTTACACAGCTCCAGGGAYGAGCTTGCAACCCYRTGGCRYACCTCTAGTAAAGAAAYT 3540  
DB 3481 GTCTTTACACAGCTCCAGGGAYGAGCTTGCAACCCYRTGGCRYACCTCTAGTAAAGAAAYT 3540  
QY 3541 GATAGTGGCAAGGGTTCRCYTCACTTTTAYGGGTAGTGGTGGCAGTGTAGTGTAGTGTAGT 3600  
DB 3541 GATAGTGGCAAGGGTTCRCYTCACTTTTAYGGGTAGTGGTGGCAGTGTAGTGTAGTGTAGT 3600  
QY 3601 GTATCTGAAGCAGTTAAATAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB 3601 GTATCTGAAGCAGTTAAATAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
QY 3661 GTGAAYRGATACCTCACTGTCTAAAGGAGCTTGTGCTCTCAGACAAACCTGTGTGTACTTAA 3720  
DB 3661 GTGAAYRGATACCTCACTGTCTAAAGGAGCTTGTGCTCTCAGACAAACCTGTGTGTACTTAA 3720  
QY 3721 TTTAGGCTCTTATTTCTGAARGGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780  
DB 3721 TTTAGGCTCTTATTTCTGAARGGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780  
QY 3781 CCAGY/CNACTTTCTTCCAGACAAATGAAGAAAGATARAAYATACTGTCTCAACAAATTAAT 3840  
DB 3781 CCAGY/CNACTTTCTTCCAGACAAATGAAGAAAGATARAAYATACTGTCTCAACAAATTAAT 3840  
QY 3841 TCTCAAAACCTAGCCACTCGAGGGACCTTGTAGAGTTCCTTGTGATGATCCYACCTT 3900  
DB 3841 TCTCAAAACCTAGCCACTCGAGGGACCTTGTAGAGTTCCTTGTGATGATCCYACCTT 3900  
QY 3901 CAACTTGTATATGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGC 3960

Db 3901 CAACCTTGATACCTGATGAGTTCCTTTGTAGAAAAAGACTTCGAAAGYGGGTATGC 3960  
Qy 3961 AGTGTGCTAGTGAATATGGAATATTTGAAAGTAATCCCTCACTCAGGAACTAGTGCTYA 4020  
Db 3961 AGTGTGCTAGTGAATATGGAATATTTGAAAGTAATCCCTCACTCAGGAACTAGTGCTYA 4020  
Qy 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACCTAGAAATTTAGGAGAAAGAAAGGYAAA 4080  
Db 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACCTAGAAATTTAGGAGAAAGAAAGGYAAA 4080  
Qy 4081 TATATATACAGACTCTRTARTATGCTYACCTAGTCNTCCATGCCCCATGMRGCAATATGSAR 4140  
Db 4081 TATATATACAGACTCTRTARTATGCTYACCTAGTCNTCCATGCCCCATGMRGCAATATGSAR 4140  
Qy 4141 AGAAGGGAATTCCTAACTTCYVAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db 4141 AGAAGGGAATTCCTAACTTCYVAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATTATYTGCGWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGTATCA 4260  
Db 4201 ATTATTATYTGCGWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCYGGGTATCA 4260  
Qy 4261 NAAAGGAAAGRAAGGGAATASAGRGAAATGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Db 4261 NAAAGGAAAGRAAGGGAATASAGRGAAATGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Qy 4321 TGCAAGGAGGACCTTCANTAGAAATGCTTAAACTTCCCTTAGTATAGGGTAATCC 4380  
Db 4321 TGCAAGGAGGACCTTCANTAGAAATGCTTAAACTTCCCTTAGTATAGGGTAATCC 4380  
Qy 4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGAGAAACAGAACTGGGNAACCTCAGG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGAGAAACAGAACTGGGNAACCTCAGG 4440  
Qy 4441 CAGTTTTCTCCCTCGGACGGTGTAGCCACTGAAAGGGAATACTTTTGCCTGCAAC 4500  
Db 4441 CAGTTTTCTCCCTCGGACGGTGTAGCCACTGAAAGGGAATACTTTTGCCTGCAAC 4500  
Qy 4501 TATCAAATGGAATTACTTAAACCCCTTCATCAAACCTTCAAACTATCAAGCARAT 4560  
Db 4501 TATCAAATGGAATTACTTAAACCCCTTCATCAAACCTTCAAACTATCAAGCARAT 4560  
Qy 4561 CCATCARATGGCAATCAATTAATTTACTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGGCAATCAATTAATTTACTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAGGCTCTGTAAGTGTGCARARAATAATCCCTGCTCTYATCGCAAGCTCTTC 4680  
Db 4621 AKTCAGGCTCTGTAAGTGTGCARARAATAATCCCTGCTCTYATCGCAAGCTCTTC 4680  
Qy 4681 AGGARAACAAARAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCACAAG 4740  
Db 4681 AGGARAACAAARAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCACAAG 4740  
Qy 4741 CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGTARATACTTTTACGGGTTGGCA 4800  
Db 4741 CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGTARATACTTTTACGGGTTGGCA 4800  
Qy 4801 RAGGCTCTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTATGAA 4860  
Db 4801 RAGGCTCTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTATGAA 4860  
Qy 4861 ATAAATCCAGATTCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTTCCAG 4920  
Db 4861 ATAAATCCAGATTCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTTCCAG 4920  
Qy 4921 GCCACGTAACCCAGGAGTATCCAGGGCTTAGGTATACGATATCACTTACACTGGCC 4980  
Db 4921 GCCACGTAACCCAGGAGTATCCAGGGCTTAGGTATACGATATCACTTACACTGGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCTCAGGAAGGTTCGAGAAAAATGAATGAAAYACTCAAAGGACATCTA 5040

Db 4981 TGAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAAYACTCAAAGGACATCTA 5040  
Qy 5041 AAAAGCAAAACCCAGGAAACCCACCTCACATGGGCTGTCTGTTCCTTATAGCCTTAAAA 5100  
Db 5041 AAAAGCAAAACCCAGGAAACCCACCTCACATGGGCTGTCTGTTCCTTATAGCCTTAAAA 5100  
Qy 5101 AGAATCTGCAACTTTTCCCAAAAAACAGGACTTAGCCCATACGAAATCTCTGTATGGAAG 5160  
Db 5101 AGAATCTGCAACTTTTCCCAAAAAACAGGACTTAGCCCATACGAAATCTCTGTATGGAAG 5160  
Qy 5161 CCCTTATAACCAATGACTTTGTCTTGAACCAAGACAGCAACTTAGTTCAGACATCA 5220  
Db 5161 CCCTTATAACCAATGACTTTGTCTTGAACCAAGACAGCAACTTAGTTCAGACATCA 5220  
Qy 5221 CCTCTTAGCCAAATATCAACAAGTTCTTAAAAACATTTACAAGGAACCTTATCCCTGAGAG 5280  
Db 5221 CCTCTTAGCCAAATATCAACAAGTTCTTAAAAACATTTACAAGGAACCTTATCCCTGAGAG 5280  
Qy 5281 AGGAAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAAT 5340  
Db 5281 AGGAAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAAT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCTGGGAAGCACTTACCCAGTCAATTTATYTTACCCCACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCTGGGAAGCACTTACCCAGTCAATTTATYTTACCCCACTG 5400  
Qy 5401 CGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATAC 5460  
Db 5401 CGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACCTTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACTCTTAGAGG 5520  
Db 5461 TGCCAAAGGAACCTTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACTCTTAGAGG 5520  
Qy 5521 ATTTGGGCTCTCTTCAAAACCAACAGGAGGAAGTAATAAAATCAATAAATCCCCC 5580  
Db 5521 ATTTGGGCTCTCTTCAAAACCAACAGGAGGAAGTAATAAAATCAATAAATCCCCC 5580  
Qy 5581 ATGSGCTCTCTTATCATATTTTCTCTKTASTGTTSTTTTACCTCTCTCTCACT 5640  
Db 5581 ATGSGCTCTCTTATCATATTTTCTCTKTASTGTTSTTTTACCTCTCTCTCACT 5640  
Qy 5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTCTATGGAGA 5700  
Db 5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTCTATGGAGA 5700  
Qy 5701 ATGCAAGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCC 5760  
Db 5701 ATGCAAGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCC 5760  
Qy 5761 ACCTTCACTGCGCAACCCCATATGCGCGCACTGCTATCACTGCGCACTCTTTGCAATG 5820  
Db 5761 ACCTTCACTGCGCAACCCCATATGCGCGCACTGCTATCACTGCGCACTCTTTGCAATG 5820  
Qy 5821 CATGCAATATCTATTATTTGGACAGGAAAAATGATTAATCTTAGTTGCTCTGGAGACT 5880  
Db 5821 CATGCAATATCTATTATTTGGACAGGAAAAATGATTAATCTTAGTTGCTCTGGAGACT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGTCTGTATGGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGTCTGTATGGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGGAAAAACATGTAAGAAAGTAATCTCCCAACTCACCGGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGGAAAAACATGTAAGAAAGTAATCTCCCAACTCACCGGGTACAT 6000  
Qy 6001 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCGTACC 6060  
Db 6001 GGCACCTCTAGCCCTCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCGTACC 6060  
Qy 6061 CATACTCGCTGGTAAGCTATTTTAATACCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 6120  
Db 6061 CATACTCGCTGGTAAGCTATTTTAATACCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 6120

6121 CAAAAACCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
Db |||||||  
6121 CAAAAACCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
Qy |||||||  
6181 CTTGTACCTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAACCACTTCCTGGTTTGTGTA 6240  
Db |||||||  
6181 CTTGTACCTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAACCACTTCCTGGTTTGTGTA 6240  
Qy |||||||  
6241 GGACCTCTTGTGTTCCAAATSTGGAATAAACCATACTCAAACTCACTGTGTGTAATAATTT 6300  
Db |||||||  
6241 GGACCTCTTGTGTTCCAAATSTGGAATAAACCATACTCAAACTCACTGTGTGTAATAATTT 6300  
Qy |||||||  
6301 AGCAATFACTACATACACCAACTCCCAATGATGATGAGTGGTAACTCTCTCCACACAA 6360  
Db |||||||  
6301 AGCAATFACTACATACACCAACTCCCAATGATGATGAGTGGTAACTCTCTCCACACAA 6360  
Qy |||||||  
6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGACCTCAGCCCTATGTTGTTTG 6420  
Db |||||||  
6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGACCTCAGCCCTATGTTGTTTG 6420  
Qy |||||||  
6421 AATGGCTCTTCAGAACTATGTCCTCTCATCTTAGTGCCCTCCATGRCATCTAC 6480  
Db |||||||  
6421 AATGGCTCTTCAGAACTATGTCCTCTCATCTTAGTGCCCTCCATGRCATCTAC 6480  
Qy |||||||  
6481 ACTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGCGCAACAAAGAGTACCCATT 6540  
Db |||||||  
6481 ACTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGCGCAACAAAGAGTACCCATT 6540  
Qy |||||||  
6541 CTTCTCTTTGTTATAGGACGAGGTGCTAGTGCTAGTGTACTAGTGTCTGGCAATTCGGGTATC 6600  
Db |||||||  
6541 CTTCTCTTTGTTATAGGACGAGGTGCTAGTGCTAGTGTACTAGTGTCTGGCAATTCGGGTATC 6600  
Qy |||||||  
6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
Db |||||||  
6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
Qy |||||||  
6661 GTCCGCGACTCCCTGTGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720  
Db |||||||  
6661 GTCCGCGACTCCCTGTGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720  
Qy |||||||  
6721 CRAAATCGAAGAGCTTTAGACTGCTAACCGCTGABAGAGGGGGACCTGTTATTTTAA 6780  
Db |||||||  
6721 CRAAATCGAAGAGCTTTAGACTGCTAACCGCTGABAGAGGGGGACCTGTTATTTTAA 6780  
Qy |||||||  
6781 GGGGAAGAATGCTGTTATTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Db |||||||  
6781 GGGGAAGAATGCTGTTATTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Qy |||||||  
6841 ATTCAGATCGAATACAACTAGCAGAGAGCTTCGAAACAACCTGGACCCCTGGGCGCTC 6900  
Db |||||||  
6841 ATTCAGATCGAATACAACTAGCAGAGAGCTTCGAAACAACCTGGACCCCTGGGCGCTC 6900  
Qy |||||||  
6901 CTCAGCCATGATGACCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATATG 6960  
Db |||||||  
6901 CTCAGCCATGATGACCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATATG 6960  
Qy |||||||  
6961 CTACTCTCTTTGGACCCCTGATCTTTTACCTCTGTTTAACTTTGCTCTCTCCAGATC 7020  
Db |||||||  
6961 CTACTCTCTTTGGACCCCTGATCTTTTACCTCTGTTTAACTTTGCTCTCTCCAGATC 7020  
Qy |||||||  
7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAAGATCTACCGCAGA 7080  
Db |||||||  
7021 GAAGCTGTRAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAAGATCTACCGCAGA 7080  
Qy |||||||  
7081 CCCCTGGACCGGCTGTTAGCCCAAGATCTGATGTTAATGACATCAAAAGGCAACCCCTCT 7140  
Db |||||||  
7081 CCCCTGGACCGGCTGTTAGCCCAAGATCTGATGTTAATGACATCAAAAGGCAACCCCTCT 7140  
Qy |||||||  
7141 GAGGAATCTAGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGAGAGTGTAGAGC 7200  
Db |||||||  
7141 GAGGAATCTAGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGAGAGTGTAGAGC 7200

7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGA 7260  
Db |||||||  
7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGA 7260  
Qy |||||||  
7261 GACAGACTAGCTGATTCCTAGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGTG 7320  
Db |||||||  
7261 GACAGACTAGCTGATTCCTAGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGTG 7320  
Qy |||||||  
7321 ACCAATCCACCTTTTAAACACCGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380  
Db |||||||  
7321 ACCAATCCACCTTTTAAACACCGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380  
Qy |||||||  
7381 CTCACATAAATGCTTAATTAGGCAACAGAGAGTAAAGAAATAGCCAATCATYATTGC 7440  
Db |||||||  
7381 CTCACATAAATGCTTAATTAGGCAACAGAGAGTAAAGAAATAGCCAATCATYATTGC 7440  
Qy |||||||  
7441 MTGAGACACAGCAGAGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCGGCA 7500  
Db |||||||  
7441 MTGAGACACAGCAGAGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCGGCA 7500  
Qy |||||||  
7501 CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Db |||||||  
7501 CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Qy |||||||  
7561 TCTATTAAATCTTGCACTGCR 7582  
Db |||||||  
7561 TCTATTAAATCTTGCACTGCR 7582

## RESULT 2

## US-10-717-580-11

; Sequence 11, Application US/10717580

; Publication No. US20040176314A1

; GENERAL INFORMATION:

; APPLICANT: BESEME, Frederic

; APPLICANT: BLOND, Jean-luc

; APPLICANT: BOUTON, Olivier

; APPLICANT: MANDRAND, Bernard

; APPLICANT: MALLET, Francois

; APPLICANT: PERRON, Herve

; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DIS-

; TITLE OF INVENTION: WITH PREGNANCY DISORDERS

; FILE REFERENCE: 105045

; CURRENT APPLICATION NUMBER: US/10/717,580

; CURRENT FILING DATE: 2003-11-21

; PRIOR APPLICATION NUMBER: US/09/446,024A

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: PCT/FR98/01442

; PRIOR FILING DATE: 1998-07-06

; PRIOR APPLICATION NUMBER: FR 97/08815

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 11

; LENGTH: 7582

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (198)..(198)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (307)..(307)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (355)..(355)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1309)..(1309)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1309)..(1309)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1331)..(1331)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (2213)..(2213)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (2398)..(2398)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (3787)..(3787)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4115)..(4115)  
 ; OTHER INFORMATION: n = any nucleotide  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4261)..(4261)  
 ; OTHER INFORMATION: n = any nucleotide

US-10-717-580-11

Query Match 100.0%; Score 7489.6; DB 9; Length 7582;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CAACAATCGGGATATAAACCAGGAGCTTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTC	60
Db	1	CAACAATCGGGATATAAACCAGGAGCTTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTC	60
Qy	61	CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTTCACTCTATTAATCTTGGCACTGCA	120
Db	61	CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGGCACTGCA	120
Qy	121	CTCTCTGCTGCATGTTTCTTACGGCTCGAGCTGAGCTTTGCTCACCGTCCACACTGC	180
Db	121	CTCTCTGCTGCATGTTTCTTACGGCTCGAGCTGAGCTTTGCTCACCGTCCACACTGC	180
Qy	181	TGTTTGGCCACACCCGACCTGCGCTGACCTCCCATCCCTCTGATCCTCGAGGCTGC	240
Db	181	TGTTTGGCCACACCCGACCTGCGCTGACCTCCCATCCCTCTGATCCTCGAGGCTGC	240
Qy	241	CGCTGTGCTTCTGATCCAGCGAGCGCCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTTG	300
Db	241	CGCTGTGCTTCTGATCCAGCGAGCGCCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTTG	300
Qy	301	CATTGTNCTGCACGGCTAAGTGCCTGGGTTTGTCTTAATGAGCTGAACACTANTCACT	360
Db	301	CATTGTNCTGCACGGCTAAGTGCCTGGGTTTGTCTTAATGAGCTGAACACTANTCACT	360
Qy	361	GGTTCCTGCTCTTCTTGTGACCCAGGCTTCTATATACTATTAACACTTACCACA	420
Db	361	GGTTCCTGCTCTTCTTGTGACCCAGGCTTCTATATACTATTAACACTTACCACA	420
Qy	421	TGGCCCAAGATTCCATTCTTGGAAATCCGCTGAGGSCAACGAACTCCAGCTCAGAGATAC	480
Db	421	TGGCCCAAGATTCCATTCTTGGAAATCCGCTGAGGSCAACGAACTCCAGCTCAGAGATAC	480
Qy	481	GARGCTTGCCACAATCTTGGAAAGCGGCTGCTACGCTTTTGGAAAGTGGTTTCAACCATC	540
Db	481	GARGCTTGCCACAATCTTGGAAAGCGGCTGCTACGCTTTTGGAAAGTGGTTTCAACCATC	540
Qy	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGGACCCMSRACGGACATCC	600
Db	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGGACCCMSRACGGACATCC	600
Qy	601	MAAGTGATGGGAAACGTTCCCGGCAAGACAAAAACGCCCTTAAGAGCTATTTCTGGARAT	660
Db	601	MAAGTGATGGGAAACGTTCCCGGCAAGACAAAAACGCCCTTAAGAGCTATTTCTGGARAT	660

Qy	661	TGGAMCAATTTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGAGTCC	720
Db	661	TGGAMCAATTTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGAGTCC	720
Qy	721	GCCTGGCAGCTCCTGAGGGAAGTATAAATTAACACCATTTACAGCTAGACYTCTTTTG	780
Db	721	GCCTGGCAGCTCCTGAGGGAAGTATAAATTAACACCATTTACAGCTAGACYTCTTTTG	780
Qy	781	TAGAAAAGGCAATGGAGTGAAGTGCCTAAGTACAACTTTCTTTTCATTAAAGACAA	840
Db	781	TAGAAAAGGCAATGGAGTGAAGTGCCTAAGTACAACTTTCTTTTCATTAAAGACAA	840
Qy	841	CTCACAATTAATGTAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC	900
Db	841	CTCACAATTAATGTAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCC	900
Qy	901	CTATCCAGCATCCCCGACTCTCTTCCCCTTAAGGACCCCTTCAACCCCAATGG	960
Db	901	CTATCCAGCATCCCCGACTCTCTTCCCCTTAAGGACCCCTTCAACCCCAATGG	960
Qy	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAAT	1020
Db	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAAT	1020
Qy	1021	TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGSCCCAGCAGAGTGCATGTCYT	1080
Db	1021	TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGSCCCAGCAGAGTGCATGTCYT	1080
Qy	1081	TTTTTCTCCAGACTTTAAAGCAATAAAAACAGACTTTAGGTAAATTTCTCAGATAAYCCT	1140
Db	1081	TTTTTCTCCAGACTTTAAAGCAATAAAAACAGACTTTAGGTAAATTTCTCAGATAAYCCT	1140
Qy	1141	GATGGCTATATGRTGTTTACAAGGGTTAGGAACAATTTCTTTGATCTGACATGGAGAT	1200
Db	1141	GATGGCTATATGRTGTTTACAAGGGTTAGGAACAATTTCTTTGATCTGACATGGAGAT	1200
Qy	1201	ATATATGCTGCTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCACCATAACTGC	1260
Db	1201	ATATATGCTGCTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCACCATAACTGC	1260
Qy	1261	AGCTGAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATGATGACAACA	1320
Db	1261	AGCTGAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATGATGACAACA	1320
Qy	1321	GAGGAAAGANAATGATTTCCCAAGCCAGCAGCTTCCAGCTTASACCTCATTTG	1380
Db	1321	GAGGAAAGANAATGATTTCCCAAGCCAGCAGCTTCCAGCTTASACCTCATTTG	1380
Qy	1381	GGGACACAGAAATCAGTAACATGGAGATTCGTCGTCAGACATTTGCTAACTTGTGTC	1440
Db	1381	GGGACACAGAAATCAGTAACATGGAGATTCGTCGTCAGACATTTGCTAACTTGTGTC	1440
Qy	1441	TASAGGACTAAGGAAACTASAGAAATCTATGAATTAATCAATGATGTCCACCAT	1500
Db	1441	TASAGGACTAAGGAAACTASAGAAATCTATGAATTAATCAATGATGTCCACCAT	1500
Qy	1501	ACACAGGGAAGGGAAGAAATCTTCTGCTTTCGAGAGACTTAAGGAGGCAATGAG	1560
Db	1501	ACACAGGGAAGGGAAGAAATCTTCTGCTTTCGAGAGACTTAAGGAGGCAATGAG	1560
Qy	1561	GAGGCTGCTCTCTGTCACCTGCTCTTCTGAGGCCCACTTAATCTTAAAGCGTAAGTT	1620
Db	1561	GAGGCTGCTCTCTGTCACCTGCTCTTCTGAGGCCCACTTAATCTTAAAGCGTAAGTT	1620
Qy	1621	TATCACTCAGTCACTGACATTTAGAAAAAACTTCAAAAAGTCTGCGTAGGCCGGAG	1680
Db	1621	TATCACTCAGTCACTGACATTTAGAAAAAACTTCAAAAAGTCTGCGTAGGCCGGAG	1680
Qy	1681	CAAACTTAGAAAAACCTTATTGAACTTTGGCAACCTTGGTATTAATAGATCAGGAG	1740
Db	1681	CAAACTTAGAAAAACCTTATTGAACTTTGGCAACCTTGGTATTAATAGATCAGGAG	1740

1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGCGCACCGCTTTAGTCATGACCT 1800  
1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGCGCACCGCTTTAGTCATGACCT 1800  
1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTGGGCAAAATGAAATGCGCTAA 1860  
1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTGGGCAAAATGAAATGCGCTAA 1860  
1861 TAGGGCTTGCTCCAGTGGGCTTACAGGACACTTTAAAAAGATTTGTCAGTAGAAG 1920  
1861 TAGGGCTTGCTCCAGTGGGCTTACAGGACACTTTAAAAAGATTTGTCAGTAGAAG 1920  
1921 TAAGCGCCCTTCGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGSCCC 1980  
1921 TAAGCGCCCTTCGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGSCCC 1980  
1981 CAGGGCAAGAGTCTTTTGAAGTCAAGAGCCACTAACCCAGATGATCCAGCAGCAGACTG 2040  
1981 CAGGGCAAGAGTCTTTTGAAGTCAAGAGCCACTAACCCAGATGATCCAGCAGCAGACTG 2040  
2041 AGGGTGGCTGGGCAAGCGCCATCCCATGCCATCAAGGACACTGAGGAGTATGCTT 2100  
2041 AGGGTGGCTGGGCAAGCGCCATCCCATGCCATCAAGGACACTGAGGAGTATGCTT 2100  
2101 GACCATTTAGGGCCAGGAGGTTGTCTCTGGACACTGTGGCGGTCTTCTTAGCTTACT 2160  
2101 GACCATTTAGGGCCAGGAGGTTGTCTCTGGACACTGTGGCGGTCTTCTTAGCTTACT 2160  
2161 CTTCTGTCCGGGCAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG 2220  
2161 CTTCTGTCCGGGCAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG 2220  
2221 GGCAGTCACTAGATATCTTTTCCAGCCACTAAGTTATGAATGAGTGGGAGCTTTATCTTT 2280  
2221 GGCAGTCACTAGATATCTTTTCCAGCCACTAAGTTATGAATGAGTGGGAGCTTTATCTTT 2280  
2281 TCATATGCTTTCTTAATATGCTTGAAGCCCACTACCTTGTGTAGGAGACATTTCTA 2340  
2281 TCATATGCTTTCTTAATATGCTTGAAGCCCACTACCTTGTGTAGGAGACATTTCTA 2340  
2341 GCAAAAGCAGGGCCATTATACCTGAACATAGGAGAGGAACACCGCTTCTGTCTNCC 2400  
2341 GCAAAAGCAGGGCCATTATACCTGAACATAGGAGAGGAACACCGCTTGTGTCTNCC 2400  
2401 CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGACAAATATGGACGAGC 2460  
2401 CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGACAAATATGGACGAGC 2460  
2461 CAAAGAAATCCCGTCTCTGTTCAAGTTAAACTAAGGATTTCCACTTCTTCCCTACCAA 2520  
2461 CAAAGAAATCCCGTCTCTGTTCAAGTTAAACTAAGGATTTCCACTTCTTCCCTACCAA 2520  
2521 GGCAGTACCCCTCAGACCCAGGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
2521 GGCAGTACCCCTCAGACCCAGGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAGTAATCCGTTAGTGGATTGAG 2640  
2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAGTAATCCGTTAGTGGATTGAG 2640  
2641 GAGGCAAGAACCCAGTGGACAGTGGAGGTTAGTGAAGATCTCAGGATTAATCAATGG 2700  
2641 GAGGCAAGAACCCAGTGGACAGTGGAGGTTAGTGAAGATCTCAGGATTAATCAATGG 2700  
2701 AGGCGGTGCTCTTTTATACCGAGTGTACCTAGCCCTTATCTGTGCTTCCCAATAC 2760  
2701 AGGCGGTGCTCTTTTATACCGAGTGTACCTAGCCCTTATCTGTGCTTCCCAATAC 2760  
2761 CAGAGGAACAGAGTGGTTTACASTCTGGACCTTGAAGATGCTTCTTCTGATCCCTG 2820  
2761 CAGAGGAACAGAGTGGTTTACASTCTGGACCTTGAAGATGCTTCTTCTGATCCCTG 2820  
2821 TACATCCTGACTCTCAATTTCTTTGTTGCTTTTGAAGATCTTTCAAAACCCARCATCTCAAC 2880

2821 TACATCCTGACTCTCAATTTCTTTGCTTTTGAAGATCTTTCAAAACCCARCATCTCAAC 2880  
2881 TCACCTGGACTTTTACCCCAAGGTTTACGGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
2881 TCACCTGGACTTTTACCCCAAGGTTTACGGGATAGYCCCCATCTATTTTGGCCAGGCAT 2940  
2941 TAGCCCAAGACTTTGAGYCAATYMTCACTGCTGACACTTTTGTCTTCTGTAAGTGGATG 3000  
2941 TAGCCCAAGACTTTGAGYCAATYMTCACTGCTGACACTTTTGTCTTCTGTAAGTGGATG 3000  
3001 ATTTTACTTTTTRGCGCYCTTCAAGAACCTTTGCGCATCAAGCCACCCAGGCTCTTMA 3060  
3001 ATTTTACTTTTTRGCGCYCTTCAAGAACCTTTGCGCATCAAGCCACCCAGGCTCTTMA 3060  
3061 ATTTTCTCGCYACTTGTGCTACAWGTTTCCAAACSAARAGCTCACTCTGCTCACAGC 3120  
3061 ATTTTCTCGCYACTTGTGCTACAWGTTTCCAAACSAARAGCTCACTCTGCTCACAGC 3120  
3121 AGTTTAAATCTTTAGGCTTAARATTAATCCAAAGGCCACCAAGGCCCTCAGTGAGGAAYRYA 3180  
3121 AGTTTAAATCTTTAGGCTTAARATTAATCCAAAGGCCACCAAGGCCCTCAGTGAGGAAYRYA 3180  
3181 TCCAGCCTATCTGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRRTTCTTCTG 3240  
3181 TCCAGCCTATCTGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRRTTCTTCTG 3240  
3241 GRTTAAAGGTTTCTCCGAAWATGATTTCCAGGTTGGGCAAAATAGCCAGYCATTTA 3300  
3241 GRTTAAAGGTTTCTCCGAAWATGATTTCCAGGTTGGGCAAAATAGCCAGYCATTTA 3300  
3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTARTAAAGATGGAAMCTGAA 3360  
3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTARTAAAGATGGAAMCTGAA 3360  
3361 GYMRAAGTGGCTTTCAGGCCCTTAAAGAGGCTTAAACCCAAAGYCCAGTGTTAAGYT 3420  
3361 GYMRAAGTGGCTTTCAGGCCCTTAAAGAGGCTTAAACCCAAAGYCCAGTGTTAAGYT 3420  
3421 TGCACACRGGGCAAGACTTTTSTTTTATVATRTACAGAAACCAACAGAAAGCTCTGGA 3480  
3421 TGCACACRGGGCAAGACTTTTSTTTTATVATRTACAGAAACCAACAGAAAGCTCTGGA 3480  
3481 GTCTTTACACAGTCCRAGGAGYAGCTTGCACACCTTTCACCTGAGTAAAGGAAAT 3540  
3481 GTCTTTACACAGTCCRAGGAGYAGCTTGCACACCTTTCACCTGAGTAAAGGAAAT 3540  
3541 GATAGTGGCAAGGTTGRCYTCATTTGTTTAYGGTATGTTGGCAGTGGCAGTYKTA 3600  
3541 GATAGTGGCAAGGTTGRCYTCATTTGTTTAYGGTATGTTGGCAGTGGCAGTYKTA 3600  
3601 GTATCTGAGCAGTTAAATATACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3601 GTATCTGAGCAGTTAAATATACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3661 GTGAAYRGCACTCACTGCTTAAAGGAGACTTGTGCTGTACAGAAACCTTGTACTTAA 3720  
3661 GTGAAYRGCACTCACTGCTTAAAGGAGACTTGTGCTGTACAGAAACCTTGTACTTAA 3720  
3721 TTTAGGCTCTATTAATTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTCTTAAC 3780  
3721 TTTAGGCTCTATTAATTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTCTTAAC 3780  
3781 CCAGY/CNCACTTTCTTCCAGAAATGAAGAAAGATARAAYATACTGTCAACAAATTAAT 3840  
3781 CCAGY/CNCACTTTCTTCCAGAAATGAAGAAAGATARAAYATACTGTCAACAAATTAAT 3840  
3841 TCTCAAACTATGCCACTCGAGGGGACCTTGTAGAGTTCCYTTGACATCCYGACCTT 3900  
3841 TCTCAAACTATGCCACTCGAGGGGACCTTGTAGAGTTCCYTTGACATCCYGACCTT 3900  
3901 CAACTTGTATCTGATGGAAGTTCTTTGTAGAAAAGGACTTTCGAAAAGYGGGTATGCG 3960

Db 3901 CAACTTGATGATGGAAGTTCCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
Qy 3961 AGTGTCAGTGATGAATGGAATYTTGAAAGTAATCCCTCCTCACTCAGGAACCTAGTGCTYA 4020  
Db 3961 AGTGTCAGTGATGAATGGAATYTTGAAAGTAATCCCTCCTCACTCAGGAACCTAGTGCTYA 4020  
Qy 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACCTAGAAATAGGAGAAAGAAAAGGYAAA 4080  
Db 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACCTAGAAATAGGAGAAAGAAAAGGYAAA 4080  
Qy 4081 TATATATACAGACTCTRTARTAGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSAR 4140  
Db 4081 TATATATACAGACTCTRTARTAGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSAR 4140  
Qy 4141 AGAAAAGGAATTCCTAACTTCYVAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db 4141 AGAAAAGGAATTCCTAACTTCYVAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATATYTGCGWGTACAGAAACCTARAGAGGTGGWAGTCTTACACTGCGGGTCACTCA 4260  
Db 4201 ATTATATYTGCGWGTACAGAAACCTARAGAGGTGGWAGTCTTACACTGCGGGTCACTCA 4260  
Qy 4261 NAAAGGAAAGAAAGGAAATASAGRGAAATGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Db 4261 NAAAGGAAAGAAAGGAAATASAGRGAAATGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Qy 4321 TGCAAGGAGGACCTCCATTTAGAAATGCTTTATTAACCTTCCTTTAGTATAGGGTAATCC 4380  
Db 4321 TGCAAGGAGGACCTCCATTTAGAAATGCTTTATTAACCTTCCTTTAGTATAGGGTAATCC 4380  
Qy 4381 CTTCCGGGNAACCAAGCCCACTACTCAGCAGAGAAACAGAACTGGGNAACCTCAGGAG 4440  
Db 4381 CTTCCGGGNAACCAAGCCCACTACTCAGCAGAGAAACAGAACTGGGNAACCTCAGGAG 4440  
Qy 4441 CAGTTTTCTCCCTCGGACGGTGTAGCCACTGAAGAGGAAATACATTTTCCCTGCAAC 4500  
Db 4441 CAGTTTTCTCCCTCGGACGGTGTAGCCACTGAAGAGGAAATACATTTTCCCTGCAAC 4500  
Qy 4501 TATCCAAATGGAATTAATTAATAACCTTCATCAAACTTTTCACTAGGCATCGATAGCAC 4560  
Db 4501 TATCCAAATGGAATTAATTAATAACCTTCATCAAACTTTTCACTAGGCATCGATAGCAC 4560  
Qy 4561 CCATCARATGGCCAAATCAATTTACTTGGAACGAGCCCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGGCCAAATCAATTTACTTGGAACGAGCCCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAGGCGCTGTGAATGTGCARARAATAATCCCTGCTCTATCGCAAGCTCCTTC 4680  
Db 4621 AKTCAGGCGCTGTGAATGTGCARARAATAATCCCTGCTCTATCGCAAGCTCCTTC 4680  
Qy 4681 AGGARAACAAARAACAGGCCATTAACCTGRARAARACTGGCAACTGATTTTACCAACAG 4740  
Db 4681 AGGARAACAAARAACAGGCCATTAACCTGRARAARACTGGCAACTGATTTTACCAACAG 4740  
Qy 4741 CCCAAACCTCAGGATTTTCACTATCTACTAGTCTGGGTARATCTTTTCAAGGTTGGGCA 4800  
Db 4741 CCCAAACCTCAGGATTTTCACTATCTACTAGTCTGGGTARATCTTTTCAAGGTTGGGCA 4800  
Qy 4801 RAGGCGCTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Db 4801 RAGGCGCTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Qy 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Db 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Qy 4921 GCCACAGTAACCCAGGAGTATCCCGAGGCTTTAGGTATACGATATCACTTACCTCGGCC 4980  
Db 4921 GCCACAGTAACCCAGGAGTATCCCGAGGCTTTAGGTATACGATATCACTTACCTCGGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCAGGAGAGTCCGAGAAATGAATGAAYATCAAAAGGACATCTA 5040  
Db 4981 TGAAGGCCACAGTCTCAGGAGAGTCCGAGAAATGAATGAAYATCAAAAGGACATCTA 5040

Qy 5041 AAAAAGCAAAACCCAGGAAAACCCACTCACATGGCTGYTCTGTGTCCTATAGCCTTAAAA 5100  
Db 5041 AAAAAGCAAAACCCAGGAAAACCCACTCACATGGCTGYTCTGTGTCCTATAGCCTTAAAA 5100  
Qy 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGG 5160  
Db 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGG 5160  
Qy 5161 CCCTTCATAACCAATGACCTTGTGCTTGACCAAGACAGCCAACTTAGTTGCGAGACATCA 5220  
Db 5161 CCCTTCATAACCAATGACCTTGTGCTTGACCAAGACAGCCAACTTAGTTGCGAGACATCA 5220  
Qy 5221 CCTCCTTAGCCAAATATCAACAAGTCTTAAAAAATTAACAAGAACCTATCCCTGAGAG 5280  
Db 5221 CCTCCTTAGCCAAATATCAACAAGTCTTAAAAAATTAACAAGAACCTATCCCTGAGAG 5280  
Qy 5281 AGGGAAGAAAGAACTATTTCCACCCWMTGACATGATTTAGTCAAGTCCCTTCCTTAATT 5340  
Db 5281 AGGGAAGAAAGAACTATTTCCACCCWMTGACATGATTTAGTCAAGTCCCTTCCTTAATT 5340  
Qy 5341 CCCCATCCTTAGATACATCTCTGGGAAGGACCTTACCAGTCAATTTATTTATACCCCACTG 5400  
Db 5341 CCCCATCCTTAGATACATCTCTGGGAAGGACCTTACCAGTCAATTTATTTATTTATACCCCACTG 5400  
Qy 5401 CGGTTAAAGTGGCTGGAGTGGATGATACATCACACTTGAGTCAAAATCCTCGATAC 5460  
Db 5401 CGGTTAAAGTGGCTGGAGTGGATGATGATACATCACACTTGAGTCAAAATCCTCGATAC 5460  
Qy 5461 TGCCAAAGAAACCTGAAATCCAGGAGACAAAGCTAGTCTTCTGTGTAACCTCTAGAGG 5520  
Db 5461 TGCCAAAGAAACCTGAAATCCAGGAGACAAAGCTAGTCTTCTGTGTAACCTCTAGAGG 5520  
Qy 5521 ATTTGCGCTCCTCTTTCAAAACAAACAGAGAGAAAGTAATAAATCATAAATCCCCC 5580  
Db 5521 ATTTGCGCTCCTCTTTCAAAACAAACAGAGAGAAAGTAATAAATCATAAATCCCCC 5580  
Qy 5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCCTTTTCACTCTCACT 5640  
Db 5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCCTTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCGCTGTATGACCAAGTACTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCCATGCGCTGTATGACCAAGTACTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGCAGCGTCCGGAAATATTTGATGCCCATCTGTATAGAGTCTTTTAAAGGGAACCC 5760  
Db 5701 ATGCAGCGTCCGGAAATATTTGATGCCCATCTGTATAGAGTCTTTTAAAGGGAACCC 5760  
Qy 5761 ACCTTCACTGCCACACCCCATATGCCCGCAACTCTATCTCTGCCACTCTTTTGCAATG 5820  
Db 5761 ACCTTCACTGCCACACCCCATATGCCCGCAACTCTATCTCTGCCACTCTTTTGCAATG 5820  
Qy 5821 CATGCAAAATCACTATTATTGGACAGAAAAATGATTAATCTCTAGTTGTCTCGAGGACTT 5880  
Db 5821 CATGCAAAATCACTATTATTGGACAGAAAAATGATTAATCTCTAGTTGTCTCGAGGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTCTTCAACCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTCTTCAACCAAACTGGTATGTCTGATGGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCNAGAGAAAACATCTAAAGAAAGTAATCTCCCACTCACCSGGGTACAT 6000  
Db 5941 CAAGATCAGGCNAGAGAAAACATCTAAAGAAAGTAATCTCCCACTCACCSGGGTACAT 6000  
Qy 6001 GGCACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 6060  
Db 6001 GGCACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 6060  
Qy 6061 CATACTCGCTGGTAAGCCTATTTTAATACACCCCTCACTGGGCTCATGAGGTCTCGGCC 6120  
Db 6061 CATACTCGCTGGTAAGCCTATTTTAATACACCCCTCACTGGGCTCATGAGGTCTCGGCC 6120



Qy	6121	CAAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCCATATGTTTCAATC	6180
Db	6121	CAAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCCATATGTTTCAATC	6180
Qy	6181	CCTGTACCTCGAACAAATGAGAAACAACCTTCAGCACAGAAATAAACACCACTCCGTTTATGTA	6240
Db	6181	CCTGTACCTCGAACAAATGAGAAACAACCTTCAGCACAGAAATAAACACCACTTCGTTTATGTA	6240
Qy	6241	GGACCTCTTGTTTCCAAATSTGGAAATAAACCCATACCTCAAACTTCACCTGTGTAAAAATTT	6300
Db	6241	GGACCTCTTGTTTCCAAATSTGGAAATAAACCCATACCTCAAACTTCACCTGTGTAAAAATTT	6300
Qy	6301	AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCGCCACAA	6360
Db	6301	AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCGCCACAA	6360
Qy	6361	ATAGTCTGCCTACCCCTCAGGAATAATTTTTTGTCTGTGTACCTCAGGCCTATCGTTGTTTG	6420
Db	6361	ATAGTCTGCCTACCCCTCAGGAATAATTTTTTGTCTGTGTACCTCAGGCCTATCGTTGTTTG	6420
Qy	6421	AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTAGTGCCCCCYATGRCCTACTAC	6480
Db	6421	AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTAGTGCCCCCYATGRCCTACTAC	6480
Qy	6481	ACTGAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAACAAAGAGTACCCATT	6540
Db	6481	ACTGAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAACAAAGAGTACCCATT	6540
Qy	6541	CTTCCTTTTGTATAGGACGAGGAGTGTAGGTGCATCTAGTACTGSCATTCGCGGTATC	6600
Db	6541	CTTCCTTTTGTATAGGACGAGGAGTGTAGGTGCATCTAGTACTGSCATTCGCGGTATC	6600
Qy	6601	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG	6660
Db	6601	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG	6660
Qy	6661	GTCCGCCACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	6720
Db	6661	GTCCGCCACTCCCTGGTCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	6720
Qy	6721	CRAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGARAGAGGGGAACTGTTTATTTTAA	6780
Db	6721	CRAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGARAGAGGGGAACTGTTTATTTTAA	6780
Qy	6781	GGGGAAGAATGCTGTTTATTTATGTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA	6840
Db	6781	GGGGAAGAATGCTGTTTATTTATGTTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA	6840
Qy	6841	ATTCSAGATCGAATACAAAGTAKAGCAGARGAGCTTCGAAACAACCTGGAACCTGGGGCCCTC	6900
Db	6841	ATTCSAGATCGAATACAAAGTAKAGCAGARGAGCTTCGAAACAACCTGGAACCTGGGGCCCTC	6900
Qy	6901	CTCAGCCRAATGGATGCCCTGGATTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATG	6960
Db	6901	CTCAGCCRAATGGATGCCCTGGATTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATG	6960
Qy	6961	CTACTCTCTTTGGACCCCTGTATCTTTTTRACCTCTCTTGTAACTTTGCTCTCTCCGAAATC	7020
Db	6961	CTACTCTCTTTGGACCCCTGTATCTTTTTRACCTCTCTTGTAACTTTGCTCTCTCCGAAATC	7020
Qy	7021	GAAGCTGTAAACTACAAATGGAGCCCAAGATCGAGTCCAAAGACTAAGATCTACCGCAGA	7080
Db	7021	GAAGCTGTAAACTACAAATGGAGCCCAAGATCGAGTCCAAAGACTAAGATCTACCGCAGA	7080
Qy	7081	CCCCTGACCGGCTGTAGCCCCACGATCTGATGTTTAAATGACATCAAAAGCACCCTCTCT	7140
Db	7081	CCCCTGACCGGCTGTAGCCCCACGATCTGATGTTTAAATGACATCAAAAGCACCCTCTCT	7140
Qy	7141	GAGGAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAAAGCAGTTAGAGC	7200
Db	7141	GAGGAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAAAGCAGTTAGAGC	7200
Qy	7201	GGTSGTCGGCCAACTCCCCCAACAGCACTTAGGTTTTTCTGTTTGAGATGGGGGACTGAGA	7260

Db	7201	GGTSGTCGGCCACCTCCCAACAGCACTTAGGTTTTCTGTGTTGAGATGGGACTGAGA	7260
Qy	7261	GACAGGACTAGCTGGAATTTCTTAGGCTGAYTAAGAATCCVTAAGCCCTAGSTGGGAAGGTG	7320
Db	7261	GACAGGACTAGCTGGAATTTCTTAGGCTGAYTAAGAATCCVTAAGCCCTAGSTGGGAAGGTG	7320
Qy	7321	ACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGYTACACCTGACCAATCAGAGAG	7380
Db	7321	ACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGYTACACCTGACCAATCAGAGAG	7380
Qy	7381	CTCACTAAATGCTAATTTAGGCAAGACAGGAGTAAGAAATAGCCAATCATYTTATTGC	7440
Db	7381	CTCACTAAATGCTAATTTAGGCAAGACAGGAGTAAGAAATAGCCAATCATYTTATTGC	7440
Qy	7441	MTGAGACACAGCAGGAGGAGCAATGATCGGATATAAACCCCAAGTTCGAGCCGGCAA	7500
Db	7441	MTGAGACACAGCAGGAGGAGCAATGATCGGATATAAACCCCAAGTTCGAGCCGGCAA	7500
Qy	7501	CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC	7560
Db	7501	CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC	7560
Qy	7561	TCTATTAAATCTTGCARCTGCR	7582
Db	7561	TCTATTAAATCTTGCARCTGCR	7582
RESULT 3			
US-09-873-367C-81			
; Sequence 81, Application US/09873367C			
; Publication No. US20030165839A1			
; GENERAL INFORMATION:			
; APPLICANT: Young, Paul			
; APPLICANT: Soppet, Daniel			
; APPLICANT: Endress, Gregory			
; APPLICANT: Augustus, Meena			
; APPLICANT: Ebner, Reinhard			
; APPLICANT: Carter, Kenneth			
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using			
; FILE REFERENCE: 689290-84			
; CURRENT APPLICATION NUMBER: US/09/873,367C			
; CURRENT FILING DATE: 2003-04-29			
; PRIOR APPLICATION NUMBER: U.S. 60/236,891			
; PRIOR FILING DATE: 2000-09-29			
; PRIOR APPLICATION NUMBER: U.S. 60/236,842			
; PRIOR FILING DATE: 2000-09-29			
; PRIOR APPLICATION NUMBER: U.S. 60/244,867			
; PRIOR FILING DATE: 2000-11-01			
; PRIOR APPLICATION NUMBER: U.S. 60/245,084			
; PRIOR FILING DATE: 2000-11-01			
; NUMBER OF SEQ ID NOS: 1067			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 81			
; LENGTH: 56093			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-873-367C-81			
Query Match 86.1%; Score 6448; DB 3; Length 56093;			
Best Local Similarity 95.7%; Pred. No. 0;			
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;			
Qy	606	GATGGGAACGTTCCCGCAAGACAAAACCGCCCTTAAGACGTATTCGGARAAATGGGA	665
Db	30953	GATGGGAACGTTCCCGCAAGACAAAACCGCCCTTAAGACGTATTCGGARAAATGGGA	31012
Qy	666	MCAATTTGACCTTCAGACACTAAGAAAGCACTTATTTCTTCGAGTGGCGCTG	725
Db	31013	CAATTTTGACCTTCAGACACTAAGAAAGCACTTATTTCTTCGAGTGGCGCTG	31072
Qy	726	GCACTCCTCGAGGAAGTATAAATATAACACCATCTTACAGCTAGACTCTTTTCTAGAA	785



Db	31073	GCACTCTCTGAGGGAAGTATAAAATTATAAACACCATCTTACAGCTAGACCTCTTTTGTAGAA	31132
Qy	786	AAGCAAAATGGAGTGAAGTGCCTAAGTACAAATCTTTTCAATTAAGAGCAACTCAC	845
Db	31133	AAGCAAAATGGAGTGAAGTGCCTAAGTACAAATCTTTTCAATTAAGAGCAACTCAC	31192
Qy	846	AATTATGTAAAAAGTGTGATTATATGCGCTACAGAGAGCCCTTCAGAGTCTACCTCCCTATC	905
Db	31193	AATTATGTAAAAAGTGTGATTATATGCGCTACAGAGAGCCCTTCAGAGTCTACCTCCCTATC	31252
Qy	906	CCAGCATCCCGCACTCTCTCCCAAMYPATAAAGGACCCCTTCACCCAAATGFTCCAA	965
Db	31253	CCAGCATCCCGCACTCTCTCCCAACTAAATAAGGACCCCTTCACCCAAATGFTCCAA	31312
Qy	966	AAGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATTCCTCCCAATATGA	1025
Db	31313	AAGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATTCCTCCCAATATGA	31372
Qy	1026	CCCTCTCCCAAGCAGTGGAGAGAGAAATTCGGCCACAGAGTGCATGTGCTTTTYY	1085
Db	31373	CCCTT-CCAGCAGTGGAGAGAGAAATTCGGCCACAGAGTGCATGTGCTTTTTC	31431
Qy	1086	TCTCCAGACTTAAAGCAATAAAAAACAGACTTAGGTAAATTCCTCAGATAAAYCCTGATGG	1145
Db	31432	TCTCCAGACTTAAAGCAATAAAAAACAGACTTAGGTAAATTCCTCAGATAAAYCCTGATGG	31491
Qy	1146	CTATATTGRTGTTTACAGGGTTAGACAAATTCCTTTGATCTGACATGGAGAGATATA	1205
Db	31492	CTATATTGATGTTTACAGGGTTAGACAAATTCCTTTGATCTGACATGGAGAGATATA-A	31550
Qy	1206	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCACCAATCTGACGCT	1265
Db	31551	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCACCAATCTGACGCT	31610
Qy	1266	GAGRGTTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGAGTANGGATCAACAAGAGG	1325
Db	31611	GAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAACAG-AGG	31667
Qy	1326	AAAGANAATGATTTCCCAACAGGCAGCAGCAGTTCCTCAGTCTASACCTTCATTTGGGAC	1385
Db	31668	AAAGAGATGATTTCCCAACAGGCAGCAGCAGTTCCTCAGTCTAGACCTTCATTTGGG---	31724
Qy	1386	ACAGAAATCAGTACATGGGAGATTGCTGTCAGACATTTGCTTAATCTGTGCTASAA	1445
Db	31725	ACACAGATCAGAACAT--GGAGATTGGTGTGTCAGACATTTGCTTAATCTGTGCTAGAA	31783
Qy	1446	GGACTTAAGGAAAATCTASGAAGAAAATCTATGAATTAATCAATGATTCACCACTAACACA	1505
Db	31784	GGACTTAAGGAAAATCTAGGAAG-AGTCTATGAATTAATCAATGATTCACCACTAACACA	31842
Qy	1506	GGGGAAGGGAAGAAAATCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTCAGGAAGC	1565
Db	31843	-GGGAAGGGAAGAAAATCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTCAGGAAGC	31901
Qy	1566	GTGCTCTCTGTCACCTGACTCTTCTGAGGCCAACTTAATCTTAAGCGTAACTTTATCA	1625
Db	31902	GTGCTCTCTGTCACCTGACTCTTCTGAGGCCAACTTAATCTTAAGCGTAACTTTATCA	31961
Qy	1626	CTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCGTAGGCCCGGAGCAAA	1684
Db	31962	CTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTCTGCGTAGGCCCGGAGCAAA	32021
Qy	1685	ACTTGAAGAACCTATTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA	1744
Db	32022	ACTTGAAGAACCTATTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA	32081
Qy	1745	AGGCGGAACAGACAAACGGGATTAATAAGAGAGCCCGCTTTAGTCATGACCCCTCAGG	1804
Db	32082	AGGCGGAACAGACAAACGGGATTAATAAGAGAGCCCGCTTTAGTCATGACCCCTCAGG	32141
Qy	1805	CAAAGTGAATTTGAGGCTCTGGAAGAGGGAAGGAAAGCTGGGCAAAATTCGAATGCTTAATAGG	1864



|||||  
35403 ATCTGCACTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 35462  
QY TTCTAAACCAATAGACTTTGTGCTTGACCCCAAGACCAACTTAGTTGCAGACATCACCT 5223  
Db TTCTAAACCAATAGACTTTGTGCTTGACCCCAAGACCAACTTAGTTGCAGACATCACCT 35522  
QY CTTTAGCCCAATATCAACAGTTCTTAAACATTTAACAACATTTAACAAGAACCTTATCCCTGAGNAGAG 5283  
Db CCTTAGCCCAATATCAACAGTTCTTAAACATTTAACAACATTTAACAAGAACCTTATCCCTGAGNAGAG 35582  
QY GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATCC 5343  
Db GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATCC 35642  
QY CATCCCTAGATACATCTCTGGGAAGACCTTACCCAGTCAATTTATPYTACCCCAACTGCGG 5403  
Db CATCCCTAGATACATCTCTGGGAAGACCTTACCCAGTCAATTTATCTACCCCAACTGCGG 35702  
QY TTAAAGTGGCTGGAGTGGAGTCTTGATACATACACTTGAGTCAAAATCTTGATACCTGC 5463  
Db TTAAAGTGGCTGGAGTGGAGTCTTGATACATACACTTGAGTCAAAATCTTGATACCTGC 35762  
QY CAAAGAACCTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT 5523  
Db CAAAGAACCTGAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT 35822  
QY TCGCGCTGCTTTCAAACAACACAGGAGGAAAGTAACTAAATCATATAATCTCCCCATG 5583  
Db TCGCGCTGCTTTCAAACAACACAGGAGGAAAGTAACTAAATCATATAATCTCCCCATG 35881  
QY GSCCTCCCTTATCATATTTCTCTKTASTGTTSTTTTACCTSTTTCACTCTCACTGCA 5643  
Db GSCCTCCCTTATCATATTTCTCTKTASTGTTSTTTTACCTSTTTCACTCTCACTGCA 35941  
QY CCCCTCCATGCGGTGTATGACCACTAGCTCCCTCTACCMAGAGTTCTATGAGAGATG 5703  
Db CCCCTCCATGCGGTGTATGACCACTAGCTCCCTCTACCMAGAGTTCTATGAGAGATG 36001  
QY CAGCGTCCCGGAAATTAATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCCACC 5763  
Db CAGCGTCCCGGAAATTAATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCCACC 36061  
QY TTCACCTGCCCAACCCATATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCCACC 5823  
Db TTCACCTGCCCAACCCATATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCCACC 36121  
QY GCAAACTACTCATTTATGGAAGGAAATGATTAATCTCTAGTTGCTTGAGGACTTGA 5883  
Db GCAAACTACTCATTTATGGAAGGAAATGATTAATCTCTAGTTGCTTGAGGACTTGA 36181  
QY GTCACTGTCTGTGGACTTACTTACCCAACTGGTATGCTGTATGCTGGGGGTGAGTTCAA 5943  
Db GTCACTGTCTGTGGACTTACTTACCCAACTGGTATGCTGTATGCTGGGGGTGAGTTCAA 36241  
QY GATCAGGCAAGGAAACCATGTAAAGAGTAATCTCCCACTCAACCSGGGTACATGGC 6003  
Db GATCAGGCAAGGAAACCATGTAAAGAGTAATCTCCCACTCAACCSGGGTACATGGC 36301  
QY ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCAT 6063  
Db ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCAT 36361  
QY ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCAT 6123  
Db ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCAT 36362  
QY AACCCCTACTAATGTTGGATATGCTCCCTGAACTTCAGGCCATATGTTTCAATCCCT 6183  
Db AACCCCTACTAATGTTGGATATGCTCCCTGAACTTCAGGCCATATGTTTCAATCCCT 36481  
QY GTACCTGAAACAACTTTCAGCACAGAAATAAACCACTTCCGTTTTAGTAGGA 6243  
Db GTACCTGAAACAACTTTCAGCACAGAAATAAACCACTTCCGTTTTAGTAGGA

Db GTACCTGAACAATGGAAACAACCTTCAGCACAGAAATAAACACCACTTCCTCGTTTTAGTAGGA 36541  
QY CCTCTGTTTTCCAATSTGGAAATAAACCATACCTCAAACTCACCCTGTGTGTAAATTTAGC 6303  
Db CCTCTGTTTTCCAATSTGGAAATAAACCATACCTCAAACTCACCCTGTGTGTAAATTTAGC 36601  
QY AATACTACATACAAACCACTCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAATA 6363  
Db AATACTACATACAAACCACTCCCAATGCAATCAGGTGGGTAACTCTCTCCACACAAATA 36661  
QY GTCTGCCCTTACCTCAGGAATATTTTTGTCTGTGGTACTCTCAGGCTATTCGTTGTGTAAT 6423  
Db GTCTGCCCTTACCTCAGGAATATTTTTGTCTGTGGTACTCTCAGGCTATTCGTTGTGTAAT 36721  
QY GGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTCATGACCATCTACACT 6483  
Db GGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTCATGACCATCTACACT 36781  
QY GAAACAAGATTTATACAGTTATGTCTAATCTAAGCCCGCAACAAAGAGTACCCATTTCTT 6543  
Db GAAACAAGATTTATACAGTTATGTCTAATCTAAGCCCGCAACAAAGAGTACCCATTTCTT 36841  
QY CTTTTTGTATAGGAGCAGGAGTCTAGGTGCATCTAGGTACTTGGCAATTTGGCGGTATCA 6603  
Db CTTTTTGTATAGGAGCAGGAGTCTAGGTGCATCTAGGTACTTGGCAATTTGGCGGTATCA 36901  
QY ACCTCTACTGATTTCTACTACAACTATCTCAAGNACTAATAATGGGACATGGAAGCGT 6663  
Db ACCTCTACTGATTTCTACTACAACTATCTCAAGNACTAATAATGGGACATGGAAGCGT 36961  
QY GCCGACTCCCTGGTCACTTGCAGAGTCAACTTAATCTCCCTAGCAGCAGTACTCTTCA 6723  
Db GCCGACTCCCTGGTCACTTGCAGAGTCAACTTAATCTCCCTAGCAGCAGTACTCTTCA 37021  
QY AATCGAAGAGCTTTAGACTTCTAAGAGTGGGAAACCTGTTATTTTAGGG 6783  
Db AATCGAAGAGCTTTAGACTTCTAAGAGTGGGAAACCTGTTATTTTAGGG 37081  
QY GAAAGATGCTGTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db GAAAGATGCTGTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 37141  
QY CSAGATCGAATACAACTGAGAGTCTGGAACACTTGAACCTTGGACCTGGGCTCTCTC 6903  
Db CSAGATCGAATACAACTGAGAGTCTGGAACACTTGAACCTTGGACCTGGGCTCTCTC 37201  
QY AGCCRAATGGATGCGCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
Db AGCCRAATGGATGCGCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 37261  
QY CTCTCTTTGGACCTGTATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATCGNA 7023  
Db CTCTCTTTGGACCTGTATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATCGNA 37321  
QY GCTGTAAACCTTACAAATGGAGCCAGATGAGTCCAAAGCTTAAGATCTACCCGAGACCC 7083  
Db GCTGTAAACCTTACAAATGGAGCCAGATGAGTCCAAAGCTTAAGATCTACCCGAGACCC 37381  
QY CTGACCCGCGCTGTAGTCTGATGTTAATGACATCAAAAGGACCCCTCTCTGAG 7143  
Db CTGACCCGCGCTGTAGTCTGATGTTAATGACATCAAAAGGACCCCTCTCTGAG 37441  
QY GAAATCTCAGCTGCAACCTCTACTAGCCCAATTTAGCGGAGCGGTAGAGCGGT 7203  
Db GAAATCTCAGCTGCAACCTCTACTAGCCCAATTTAGCGGAGCGGTAGAGCGGT 37501  
QY SSTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGTGGGGACTTGAGAGAC 7263  
Db SSTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGTGGGGACTTGAGAGAC 37561  
QY AGGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCYTAAGCCTTAGSTGGGAGGTGACC 7323  
Db AGGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCYTAAGCCTTAGSTGGGAGGTGACC 37621

QY	7324	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCACACCTTGACCAATCAGAGAGCTC	7383
Db	37622	ACATCCACCTTTAAACACACGGGGCTTGCAACTTAGCTCACAACCTGACCAATCAGAGAGCTC	37681
QY	7384	ACTAAATGCTAATTAGGCAAAACAGCAGGAGGTAAACAAATAGCCATCATYTTATTGCMTG	7443
Db	37682	ACTAAATGCTAATTAGGCAAAACAGCAGGAGGTAAACAAATAGCCATCATCTATTGSCCTG	37741
QY	7444	AGACACAGCAGGAGGAGCAATGATCGGATATAAACCCCAAGYTTTCAGACCGGCAACGG	7503
Db	37742	AGACACAGCAGGAGGAGCAATGATCGGATATAAACCCCAAGCTTCGAGCGCGCAACGG	37801
QY	7504	CAACCCCTTTGGTCCCTCCCTTTGTATGAGGAGCTCTGTTTTCATGCTATTTCACCTCT	7563
Db	37802	CAACCCCTTTGGTCCCTCCCTTTGTATGAGGAGCTCTGTTTTCATGCTATTTCACCTCT	37861
QY	7564	ATTAATCTTGCARCTGCR	7582
Db	37862	ATTAATCTTGCARCTGCA	37880
RESULT 4			
US-10-843-641A-81			
; Sequence 81, Application US/10843641A			
; Publication No. US20050064454A1			
; GENERAL INFORMATION:			
; APPLICANT: Avalon Pharmaceuticals, Inc.			
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using			
; TITLE OF INVENTION: Signature Gene Sets			
; FILE REFERENCE: 689290-189			
; CURRENT APPLICATION NUMBER: US/10/843,641A			
; PRIOR FILING DATE: 2004-05-12			
; PRIOR APPLICATION NUMBER: US/09/873,367			
; PRIOR FILING DATE: 2001-06-05			
; PRIOR APPLICATION NUMBER: US/09/954,531			
; PRIOR FILING DATE: 2001-09-18			
; PRIOR APPLICATION NUMBER: US/09/954,456			
; PRIOR FILING DATE: 2001-09-25			
; PRIOR APPLICATION NUMBER: US/09/962,436			
; PRIOR FILING DATE: 2001-09-25			
; PRIOR APPLICATION NUMBER: US/09/962,832			
; PRIOR FILING DATE: 2001-09-25			
; PRIOR APPLICATION NUMBER: US/09/964,824			
; PRIOR FILING DATE: 2001-09-27			
; PRIOR APPLICATION NUMBER: US/09/967,768			
; PRIOR FILING DATE: 2001-09-28			
; PRIOR APPLICATION NUMBER: US/09/968,007			
; PRIOR FILING DATE: 2001-10-02			
; PRIOR APPLICATION NUMBER: US/09/969,347			
; PRIOR FILING DATE: 2001-10-02			
; PRIOR APPLICATION NUMBER: US/09/969,708			
; PRIOR FILING DATE: 2001-10-03			
; Remaining prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 8447			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 81			
; LENGTH: 56093			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-843-641A-81			

QY	726	GC	ACTCTCTCAGGGAGTATAAAATTATAACCACTTTACAGCTAGAC	YCTCTTTTGTAGAA	785
Db	31073	GC	ACTCTCTCAGGGAGTATAAAATTATAACCACTTTACAGCTAGAC	CTCTTTTGTAGAA	31132
QY	786	RA	GGCAAAATGGAGTGAAGTGCCATAGTACAAACTTCTTTTCATTAAGAGACAACTCAC	845	
Db	31133	RA	GGCAAAATGGAGTGAAGTGCCATAGTACAAACTTCTTTTCATTAAGAGACAACTCAC	31192	
QY	846	AA	TATGTAAAAAGTGTGATTTATGCGCTACAGAAAGCCCTTCAGAGTCTACCTCCCTATC	905	
Db	31193	AA	TATGTAAAAAGTGTGATTTATGCGCTACAGAAAGCCCTTCAGAGTCTACCTCCCTATC	31252	
QY	906	CG	AGCATCCCCGACTCTTCTCCCACTAAATTAAGGACCCCTTCAACCCAAATGGTCCAA	965	
Db	31253	CG	AGCATCCCCGACTCTTCTCCCACTAAATTAAGGACCCCTTCAACCCAAATGGTCCAA	31312	
QY	966	AA	GGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTCCTCCCAATTAATGA	1025	
Db	31313	AA	GGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTCCTCCCAATTAATGA	31372	
QY	1026	CC	CTCTCCCAAGCAGTGGGAGGAGAGAAATTCGCGCCAGCCAGAGTGATGTCGTCVTTTTY	1085	
Db	31373	CC	CTCTCTCCCAAGCAGTGGGAGGAGAGAAATTCGCGCCAGCCAGAGTGATGTCGTCVTTTTY	31431	
QY	1086	TC	TCCAGACTTAAAGCAATAAAAAACAGACTTAGGTAAATCTCAGATAAATCTGATGG	1145	
Db	31432	TC	TCCAGACTTAAAGCAATAAAAAACAGACTTAGGTAAATCTCAGATAAATCTGATGG	31491	
QY	1146	CT	ATATTTGTTGTTTTACAAGGGTTAGGACAAATCTCTTTTGATCTGACATGGAGAGATATA	1205	
Db	31492	CT	ATATTTGTTGTTTTACAAGGGTTAGGACAAATCTCTTTTGATCTGACATGGAGAGATATA	31550	
QY	1206	TG	CTACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCCACCATTAACCTGCAGCCT	1265	
Db	31551	TG	CTACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCCACCATTAACCTGCAGCCT	31610	
QY	1266	GAG	GTGTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACACAGAAAG	1325	
Db	31611	GAG	GTGTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACACAGAAAG	31667	
QY	1326	AA	AGANAATGATTTCCCAAGCCAGCARGCAGTTCCAGTCTPASACCTCATTTGGGGAC	1385	
Db	31668	AA	AGAGAATGATTTCCCAAGCCAGCARGCAGTTCCAGTCTPASACCTCATTTGGGGAC	31724	
QY	1386	AC	AGAAATCAGTAACATGGGAGATTGGTCTCGCAGACATTTGCTAACTGTGTGCTASAA	1445	
Db	31725	AC	AGAAATCAGTAACATGGGAGATTGGTCTCGCAGACATTTGCTAACTGTGTGCTASAA	31783	
QY	1446	GG	ACTTAAGGAAATCTASGAAGAAATCTAYGAATTAATCAATGATCTCCACCATTAACACA	1505	
Db	31784	GG	ACTTAAGGAAATCTASGAAGAAATCTAYGAATTAATCAATGATCTCCACCATTAACACA	31842	
QY	1506	GG	GGAAGGGAAGAAATCTACTGCTTTCTCGAGAGACTAAGGGAGGCAATTTGAGGAAGC	1565	
Db	31843	GG	GGAAGGGAAGAAATCTACTGCTTTCTCGAGAGACTAAGGGAGGCAATTTGAGGAAGC	31901	
QY	1566	GT	GCCTCTCTGCTCACTGACTCTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTTATCA	1625	
Db	31902	GT	GCCTCTCTGCTCACTGACTCTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTTATCA	31961	
QY	1626	CT	CAGTCTGAGTGCAGACATTAG-AAAAACTTTCAAAGTCTCCGCTAGGCCCGGAGC	1684	
Db	31962	CT	CAGTCTGAGTGCAGACATTAG-AAAAACTTTCAAAGTCTCCGCTAGGCCCGGAGC	32021	
QY	1685	ACT	TAGAAACCTATTGTAACCTTGGCAACTTCGGTTTTTTTTTATTAATAGAGATCAGGAGGAGC	1744	
Db	32022	ACT	TAGAAACCTATTGTAACCTTGGCAACTTCGGTTTTTTTTTATTAATAGAGATCAGGAGGAGC	32081	
QY	1745	AG	GCGGAAACAGACAAACGGGATTAATAAAAAAGGCCACCGCTTTTATGTCATGACCTTCAGG	1804	
Db	32082	AG	GCGGAAACAGACAAACGGGATTAATAAAAAAGGCCACCGCTTTTATGTCATGACCTTCAGG	32141	

Qy	1805	CAAGTGCATTTGGAGGCTCTGAAAGGAAAGCAAGCTGGGCAAAATTGAATGCCCTAAATAGG	1864
Db	32142	CAAGTGCATTTTGGAGGCTCTGAAAGGAAAGCAAGCTGGGCAAAATTGAATGCCCTAAATAGG	32201
Qy	1865	GCATTGCTTCCAGTGCGGTCTACAAGGACACTTTTAAAAAGATTGTGCCAAGTAGAAGTAAG	1924
Db	32202	GCATTGCTTCCAGTGCGGTCTACAAGGACACTTTTAAAAAGATTGTGCCAAGTAGAAGTAAG	32261
Qy	1925	CCGCCCTTTCGTCCATGCCCCCTTATTTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG	1984
Db	32262	CCGCCCTTTCGTCCATGCCCCCTTATTTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG	32321
Qy	1985	GGCAAAAGTGTTTTTGAGTCAGAAGCCAATAACAGATGATCCAGACGAGCACTGAGGG	2044
Db	32322	GGCAAAAGTGCTCTGAGTCAGAAGCCAATAACAGATGATCCAGACGAGCACTGAGGG	32381
Qy	2045	TGCTTGGGGCAAGGCCATCCATGCCATCACCTTCACAGAGCCCTGGGTATGTTGACC	2104
Db	32382	TGCTTGGGGCAAGGCCATCCATGCCATCACCTTCACAGAGCCCTGGGTATGTTGACC	32441
Qy	2105	ATTGAGGGCCAGGAAGTTGTCTCTCGGACACTGGTGGCGTCTTCTTAGTCTTACTCTTC	2164
Db	32442	ATTGAGGGCCAGG-AGTTGTCTCTCGGACACTGGTGGCGTCTTCTTAGTCTTACTCTTC	32500
Qy	2165	TGTCGCCGCAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGGCA	2224
Db	32501	TGTCGCCGCAACTGTCCTCCAGATCTGTCACTA-TCTGAGGGGGTCC- TAAGACGGGCA	32558
Qy	2225	GTCACTAGATACTTTTTCCAGGCCACTAAGTTATGAACHTGGGGAGCTTTATCTTTTAC	2284
Db	32559	GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG- ACTGGGGAGCTTTATCTTTTTCAC	32616
Qy	2285	ATGCTTTTCTAAATTATGCTTGAAGGCCCACTACCTTGTTTAGGAGAGACATTTTAGCAA	2344
Db	32617	ATGCTTTTCTAAATTATGCTTGAAGGCCCACTACCTTGTTTAGGAGAGACATTTTAGCAA	32676
Qy	2345	AAGCAGGGGCCATTATACACTGAACATAGGAGAGGAACAACCCGTTTGTGTGTTGCCCTG	2404
Db	32677	AAGCAGGGGCCATTATAGACCTGAACATAGGAGAGGAACAACCCGTTTGTGTGTTGCCCTG	32735
Qy	2405	CTTGAGGAAGGAATTAATCCTGAACTCTGGGCACAGAAGGACAATATGGAGGCCAA	2464
Db	32736	CTTGAGGAAGGAATTAATCCTGAACTCTGGGCACAGAAGGACAATATGGAGCGAG-CAAA	32794
Qy	2465	GAATGCCGCTCTGTTTCAAGTTAACTAAAGGANTCCACTCTCTTCCCTACCAAGGCA	2524
Db	32795	GAATGCCGCTCTGTTTCAAGTTAACTAAAGGANTCCACTCTCTTCCCTACCAAGGCA	32854
Qy	2525	GTACCCCTCTCAGACCCCAAGGCCCAACAGGATTTCCAAAAGATTGTTAAGGACTTAAAGC	2584
Db	32855	GTACCCCTCTCAGACCCCAAGGCCCAACAGGACTTCCAAAAGATTGTTAAGGACTTAAAGC	32914
Qy	2585	CCAAAGGCTTAGTAAACCATGCAATACTCCTCGAGTAATTCGTAGTGGATTTAGGAGG	2644
Db	32915	CCAAAGGCTTAGTAAACCATGCAAGTAAACCCCTGCAGTACTCC-----AATTTTAGGAG	32967
Qy	2645	CACAGAAACCCAGTGGACAGCTGGAGGGTAGTGCAGATCTCAGGATTTATCAATGGAGC	2704
Db	32968	TACAGAAACCCCAACAGACAGTGA- GGTAGTGCAGATCTCAGGATTTATCAAT- GAGGC	33025
Qy	2705	CGTTGCTCTTTTATACCACCGCTGTACTAGCCCTTATATCTGTGMYTTCCTCAATACACAGA	2764
Db	33026	TGTTGTTCTCTATAGCCAGCTGTACTAGCCCTTATATCTGTGCTTTCCANATACCAGA	33085
Qy	2765	GGAAACAGAGTGGTTTACSTCTCTGGAACCTTMAAGATGCTTCTTCTGCAATCCCTGTACA	2824
Db	33086	GGAAACAGAGTGGTTTACAGTCTCTGGAACCTTCAAGGATGCTTCTTCTGCAATCCCTGTACA	33145
Qy	2825	TCCTGACTCTCAATTTCTGTTTGCCTTGAAGATCTTCAAAACCCACATCTCAACTCAC	2884
Db	33146	TCCTGACTCTCAATTTCTGTTTGCCTTGAAGATCTTCAAAACCCACATCTCAACTCAC	33205
Qy	2885	CTGACTCTTTTTTACCCCAAGGGTTACGGGATAGYCCCCCATCTATTTTGGCCAGGCAATTAGC	2944

Db	33206	CTGACATATTTTACCCCAAGGGTTACAGGATAGTCCCATCTATTTTGGCCAGCATTTAGC	33265
Qy	2945	CAAAGACTTGAGYCARVTYMTCATACCTGGACACTCTTGTCTTTCRGTAAGTGGATGATTT	3004
Db	33266	CCAAGACTTGAGCCAACTCTCATACCTGGACA--CTTGTCTCTCGTAGGTGGATGATTT	33323
Qy	3005	ACTTTTTRGCGCCYRTRTTCAGAAACCTTGTGCATCAAGCCACCCCAAGCRCTCTTAAATTT	3064
Db	33324	ACTTTTGGCGCCCATTCAGAAACCTTGTGCATCAAGCCACCCCAAGCGCTCTTCAATTT	33383
Qy	3065	CCTGCTACCTGTGGCTACAGGTTTCCAAACGAPARGCTCARCTCTGCTCACAGCAGT	3124
Db	33384	CCTGCTACCTGTGGCTACATGGTTTCCAAACCAAGGCTCAACTCTGCTCACAGCAGT	33443
Qy	3125	TAAATACCTTAGRCCTTAARATTTATCCAAAGGACCAARGGCCCTCAGCTGAGAAAYRVATCCA	3184
Db	33444	----TACTTAGGCTTAARATTTATCCAAAGGACCAAGGCCCTCAGTGAGAAACCAATCCA	33499
Qy	3185	GCCTATCTGGCTTATCCTCATCYAAAAACCTTAAAGCAACTAAGRGRRRTTCTTTGGCRT	3244
Db	33500	GCCTATCTGGCTTATCCTCATCCCAAAACCTTAAAGCAACTAAGGGGATTCCTTTGGCGT	33559
Qy	3245	AAYAGGYTTCGCCAAAWATGATTCGCCAGGTWTGGCBAATATAGCCAGGYCATTAWATTA	3304
Db	33560	AATAGGTTTCTGCCAAATGGAAT-CCAGGTATGGCGAAATAGCCAGGTCAATTAATA	33618
Qy	3305	CASATAATTAAGGAACTCAGAAAGCCAAATACCCATTTTARTAGATGGAYAMCTGAAGYMR	3364
Db	33619	CACATAATTAAGGAACTCAGAAAGCCAAATACCCATTTTARTAGATGGACACTGAAGTAG	33678
Qy	3365	AAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCAGTGTTAAGYTTGCC	3424
Db	33679	AAGTGGCTTTCAGGCC-----TAACCCCAAGCCCCAGTGTTAAGTTTTGGC	33724
Qy	3425	AACRGGGCAAGACTTTTSTTATYVRTCACAGAAAAACAGRAAYAGCTCTRGGAGTCC	3484
Db	33725	AACAGGGCAAGACTTTTCTCATATGTCCACAG-AAAAAACAGGAATAGCTCTAGGAGTCC	33783
Qy	3485	TTACACAGTCCRAGGGAYGAGTTGCCAACCVRTGGRYACCTTGASTAGGAAAYTGATG	3544
Db	33784	TTACACAGTCCGAGGGATGAGCTTGCAACCTGTGGCATACCTGACTAAGGAAYTGATG	33843
Qy	3545	TAGTGGCAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTAT	3604
Db	33844	TAGTGGCAAGGGTTGACCTCATTTGTTAGGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT	33903
Qy	3605	CTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAGTGA	3664
Db	33904	CTGAAGCAGTTTAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGA	33963
Qy	3665	AVRGCATACTCACTCTTAAAGGAGACTTGTGGCTGTCCAGACAACVGTTTACTTTAAATRTC	3724
Db	33964	ATGGCATACTCACTCTTAAAGGAGACTTGTGGCTGTCCAGACAACVGTTTACTTTAAATGTC	34023
Qy	3725	AGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACCTTGTGCAACTCTTTAAACCCAG	3784
Db	34024	AGGCTCTATTACTTGAAGGGCAGTGTGCGACTGTGCACTTGTGCAACTCTTTAAACCCAG	34083
Qy	3785	YCNCAATTTCTCCAGACAATGAAGAAAAAGATAPAAAYATAACTGTCCAACAARTAAATTTCTC	3844
Db	34084	CCACATTTCTCCAGACAATGAAGAAAAAGATAAAACATAAATGTCCAACAAGTAATTTCTC	34143
Qy	3845	AAACCTATGCCACTCGAGGGGACCTTATAGAGTTCCTTGTGACTGATCCVGCACCTTCAAC	3904
Db	34144	AAACCTATGCCACTCGAGGGGACCTTTTATAGAGTTCCTTGTGACTGATCCCGACC--TCAAC	34202
Qy	3905	TTGTATCTATGATGGAAGTTCTTCTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGCAGTG	3964
Db	34203	TTGTATCTATGATGGAAGTTCTTCTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGCAGTG	34262
Qy	3965	GTCACTGATTAATGGAATAYTTGAAAGTAAATCCCTTCACCTCCAGGAACCTAGTGTCTYAGCTR	4024





QY 6184 GTACCTGAACAAATGGAACAACTTCAGCACAGAAAAATAAACACACACTTCGTTTTAGTAGGA 6243  
 Db 36482 GTACCTGAACAAATGGAACAACTTCAGCACAGAAAAATAAACACACACTTCGTTTTAGTAGGA 36541  
 QY 6244 CCTCTGTTTCCAAATSTGGAATAAACCACATACCTCAAACTCAGCTGTGTAAATTTAGC 6303  
 Db 36542 CCTCTGTTTCCAAATSTGGAATAAACCACATACCTCAAACTCAGCTGTGTAAATTTAGC 36601  
 QY 6304 AATACATACATACACAAACCACTCCCAATGATCAGGTGGGTAACTCCTCCCAACAAATA 6363  
 Db 36602 AATACATACATACACAAACCACTCCCAATGATCAGGTGGGTAACTCCTCCCAACAAATA 36661  
 QY 6364 GTCTGCTACCTCAGGAATAATTTTGTCTGTGTACCTCAGCTATGCTGTGTTTGAAT 6423  
 Db 36662 GTCTGCTACCTCAGGAATAATTTTGTCTGTGTACCTCAGCTATGCTGTGTTTGAAT 36721  
 QY 6424 GGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTTAGTGCCCCYATGRCATCTACACT 6483  
 Db 36722 GGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTTAGTGCCCCYATGRCATCTACACT 36781  
 QY 6484 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGGCAACAAAGAGTACCCATTTCT 6543  
 Db 36782 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGGCAACAAAGAGTACCCATTTCT 36841  
 QY 6544 CCTTTGTTTATAGGAGCAGGAGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCACA 6603  
 Db 36842 CCTTTGTTTATAGGAGCAGGAGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCACA 36901  
 QY 6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 6663  
 Db 36902 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 36961  
 QY 6664 GCGCACTCCTGGTCACCTTGAAGATCACTTAACCTCCTAGCAGCAGTAGTCTCTTCA 6723  
 Db 36962 GCGCACTCCTGGTCACCTTGAAGATCACTTAACCTCCTAGCAGCAGTAGTCTCTTCA 37021  
 QY 6724 AATCGAAGCTTTAGACTGCTTAACCGCTGAGAGGGGGACCTGTTTATTTTAGG 6783  
 Db 37022 AATCGAAGCTTTAGACTGCTTAACCGCTGAGAGGGGGACCTGTTTATTTTAGG 37081  
 QY 6784 GAAGAATGCTGTTATTTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTAAGAAAT 6843  
 Db 37082 GAAGAATGCTGTTATTTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTAAGAAAT 37141  
 QY 6844 CSAGATCGAATAACAAGTAKAGCAGARGACTTCGAAAACACTGGACCTGGGGCCTCTC 6903  
 Db 37142 CGAGATCGAATAACAAGTAKAGCAGARGACTTCGAAAACACTGGACCTGGGGCCTCTC 37201  
 QY 6904 AGCCATGATGCTGCTGGAATCTCCCTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
 Db 37202 AGCCATGATGCTGCTGGAATCTCCCTCTTAGGACCTCTAGCAGCTATATATTTGCTA 37261  
 QY 6964 CTCCTCTTTGGACCTCTGATCTTTACCTCCTTTGTTAACTTTGCTCTCTCCAGAAATCGAA 7023  
 Db 37262 CTCCTCTTTGGACCTCTGATCTTTACCTCCTTTGTTAACTTTGCTCTCTCCAGAAATCGAA 37321  
 QY 7024 GCTGTAAACTACAAATGAGGCCCAAGATCGAGTCCAGACTAAGATCTACCCGAGACC 7083  
 Db 37322 GCTGTAAACTACAAATGAGGCCCAAGATCGAGTCCAGACTAAGATCTACCCGAGACC 37381  
 QY 7084 CTGGACCGGCTGTAGCCCAAGATCTGATGTTAATGACATCAAGGCGACCCCTCTGAG 7143  
 Db 37382 CTGGACCGGCTGTAGCCCAAGATCTGATGTTAATGACATCAAGGCGACCCCTCTGAG 37441  
 QY 7144 GAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAAGAGCTTAGAGCGGT 7203  
 Db 37442 GAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAAGAGCTTAGAGCGGT 37501  
 QY 7204 SGTGCGCCAACTCCCAACAGACCTTAGGTTTCTGTTGATGGGGAGCTGAGAGAC 7263  
 Db 37502 SGTGCGCCAACTCCCAACAGACCTTAGGTTTCTGTTGATGGGGAGCTGAGAGAC 37561  
 QY 7264 AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGSTGGAGGTGACC 7323

Db 37562 AGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAGGTGACC 37621  
 QY 7324 ACATCCACCTTTAAACACGGGCTTGCNACTTAGTGTACACCTGACCAATCAGAGAGCTC 7383  
 Db 37622 ACATCCACCTTTAAACACGGGCTTGCNACTTAGTGTACACCTGACCAATCAGAGAGCTC 37681  
 QY 7384 ACTAAATGCTTAATTAGGCAAGACAGGAGTAAAGAAATAGCCAATCATYATTGOWTG 7443  
 Db 37682 ACTAAATGCTTAATTAGGCAAGACAGGAGTAAAGAAATAGCCAATCATYATTGOWTG 37741  
 QY 7444 AGAGCACAGCAGGAGGACCAATGATCGGATATATAACCAAGTATTTCAGCGCGCAACGG 7503  
 Db 37742 AGAGCACAGCAGGAGGACCAATGATCGGATATATAACCAAGTATTTCAGCGCGCAACGG 37801  
 QY 7504 CAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563  
 Db 37802 CAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTTCATGCTATTTCACTCT 37861  
 QY 7564 ATTAAATCTTCGACTGCR 7582  
 Db 37862 ATTAAATCTTCGAACTGCA 37880

RESULT 5

US-11-028-539-3  
 ; Sequence 3, Application US/11028539  
 ; Publication No. US20050118573A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PERIN, Jean-Pierre  
 ; APPLICANT: RIEGER, Francois  
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU  
 ; FILE REFERENCE: 200936USOPT  
 ; CURRENT APPLICATION NUMBER: US/11/028,539  
 ; CURRENT FILING DATE: 2005-01-05  
 ; PRIOR APPLICATION NUMBER: US/09/719,554  
 ; PRIOR FILING DATE: 2001-01-18  
 ; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
 ; PRIOR FILING DATE: 1999-06-25  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 10499  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-11-028-539-3

Query Match 85.9%; Score 6436; DB 13; Length 10499;  
 Best Local Similarity 95.7%; Pred. No. 0;  
 Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

QY 606 GATGGGAACGTTCCCGCAAGACAAAAACGCCCTTAAGACGTATTTGGAAAAATTTGGA 665  
 Db 2953 GATGGGAACGTTCCCGCAAGACAAAAACGCCCTTAAGACGTATTTGGAAAAATTTGGA 3012  
 QY 666 MCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCCGCTG 725  
 Db 3013 CCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCCGCTG 3072  
 QY 726 GCATCTCTGAGGAAGTAAATTAATAACCATCTTACAGCTAGACACTCTTTTGTAGAA 785  
 Db 3073 GCATCTCTGAGGAAGTAAATTAATAACCATCTTACAGCTAGACACTCTTTTGTAGAA 3132  
 QY 786 AAGCAAAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGAGACAACCTAC 845  
 Db 3133 AAGCAAAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGAGACAACCTAC 3192  
 QY 846 AATATGTAAGAAAGTGAATTTATGCCCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC 905  
 Db 3193 AATATGTAAGAAAGTGAATTTATGCCCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC 3252





Db 5384 CCTCGCTACCTGTGGCTATCGTTTCCAAACCAAGGCTCAACTCTGCTCACAGCAGGT 5443  
Qy 3125 TAAATACCTTAGRCCTAARATTTATCAAAAGGCACCAAGGCCCTCAGTGGAGAAATATCA 3184  
Db 5444 ---TACTTAGGGCTAAATATTCANAAAGGCHCAGGGCCCTCAGTGAGGAACATCCCA 5499  
Qy 3185 GCCTATACCTGGCTTATCTCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTTGGGCT 3244  
Db 5500 GCCTATACCTGGCTTATCTCTCATCCCAAAACCCCTAAGCAACTAAGGGGATTCCTTGGGCT 5559  
Qy 3245 AAYAGGYTTCTGCCAANATGATTTCCAGAGTWTGGCRAAATAGCCAGGCAATTAATATA 3304  
Db 5560 AATAGGTTTCTGCCGAAATGGATT-CCCAAGTATGGGCAATAGCCAGGTCATTAATA 5618  
Qy 3305 CASTAATTAAGGAACCTCAGAAAGCAATACCCATTTARTAGATGGAYMCTGAAGYMR 3364  
Db 5619 CACTAATTAAGGAACCTCAGAAAGCAATACCCATTTARTAGATGGCAACTGAAGTAG 5678  
Qy 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAAGTGTAAAGYTTGCC 3424  
Db 5679 AAGTGGCTTTCCAGGCC-----TAAACCCCAAGCCCAAGTGTAAAGYTTGCC 5724  
Qy 3425 AACGGGCAAGACTTTTSTYATAYRTCAAGAAAAAAACAGRAAYAGCTCTRGAGTCC 3484  
Db 5725 AACGGGCAAGACTTTTCTCATATGTGTACAG-AAAAAACAGGAATAGCTCTAGGAGTCC 5783  
Qy 3485 TTACACAGTTCBAGGAYGACCTTGCACACCVRTGGCRYACCTGASTAAGGAATGTATG 3544  
Db 5784 TTACACAGATCCGAGGGATGAGCTTGCACACCTGTGGCATACCTGACTAAGGAATGTATG 5843  
Qy 3545 TAGTGGCAAGGTTGRCVYTCATTTAYGGGTAGTGTGGCAGTACGAGYKTAGTAT 3604  
Db 5844 TAGTGGCAAGGTTGACCTCAITGTTACGGGTAGTGTGGCAGTACGAGCTTAGTAT 5903  
Qy 3605 CTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
Db 5904 CTGAAGCAGTTAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGA 5963  
Qy 3665 AYRGACTACTGCTAAGAGAGACTTGTGGCTGTACAGAAACVGTTTACTTAATRTC 3724  
Db 5964 ATGGCATACTCACTGCTAAGAGAGACTTGTGGCTGTACAGAAACVGTTTACTTAATGTC 6023  
Qy 3725 AGGCTCTATTACTTGAAGGGCCAGTGTGCTGACTGTGCACTTGTGCACTCTTAACCCAG 3784  
Db 6024 AGGCTCTATTACTTGAAGGGCCAGTGTGCTGCACTGTGCACTTGTGCACTCTTAACCCAG 6083  
Qy 3785 YNCATTTCTCCAGACATGAAGAAAGATARAAYATACTGTCAACAAATPAATTTCTC 3844  
Db 6084 CCACATTTCTCCAGACATGAAGAAAGATARAATAAATAAATAAATAAATAAATAAATA 6143  
Qy 3845 AAACCTATGCCACTCGAGGGGACCTTGTAGAGTTCCTTGTGACTGATCCYGCCTTCAAC 3904  
Db 6144 AAACCTATGCCACTCGAGGGGACCTTGTAGAGTTCCTTGTGACTGATCCCGACC-TCAAC 6202  
Qy 3905 TTGTATACTGATGGAGTTCCTTGTAGAAAGGACTTCGAAAGYGGGGTATCGATG 3964  
Db 6203 TTGTATACTGATGGAGTTCCTTGTAGAAAGGACTTCGAAAGYGGGGTATCGATG 6262  
Qy 3965 GTCAGTGAATAAGGAATVTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTVAGCTR 4024  
Db 6263 GTCAGTGAATAAGGAATVTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTVAGCTR 6322  
Qy 4025 GCAGAACTAATAGCCYTCATYTKGGGCACTAGAAATTAGGAGAGRAAAAGGGYAAATATA 4084  
Db 6323 GCAGAACTAATAGCCCTCACTTTGGGCACTAGAAATTAGGAGAGRAAAAGGGYAAATATA 6382  
Qy 4085 TATACAGACTCTEARTATCTYACCTAGTCNTCCATGCCCCATGMBGCATATGSRAGAA 4144  
Db 6383 TATACAGACTCTEARTATCTYACCTAGTCNTCCATGCCCCATGCCCCATATGSRAGAA 6442  
Qy 4145 AGGGAATTCCTAACTTCYAGAGGAAACCTATCAMACATCAGGAAGCCATTAGGARATTA 4204  
Db 6443 AGGGAATTCCTAACTTCYAGAGGAAACCTATCAMACATCAGGAAGCCATTAGGARATTA 6502

Qy 4205 TTAYTGGCWGTACAGAAACCTARAGAGTGTGMAGTCTTTACACTGCTGCGGGTCAATCANAAA 4264  
Db 6503 TTATTTGGCTGTACAGAAACCTAAGAGGTGGCAGTCTTACACTGCGGGGTCAATCAGAAA 6562  
Qy 4265 GGAAGRAAAGGGGAAATASAGGAAATGTCAGCAACATATTTGAAGCMAAAGAGCTGCA 4324  
Db 6563 GGAAGRAAAGGGGAAATAGAGAGAACTGCGCAAGCAGATATTTGAAGCCAAAAGAGCTGCA 6622  
Qy 4325 AGGAGGAGCCCTCCATTTAGAAATGCTTTAAACCTTCCCTTTAGTATAGGTAATCCCTTC 4384  
Db 6623 AGGAGGAGCCCTCCATTTAGAAATGCTTTAAACCAACCCCTTAGTATAGGTAATCCCTTC 6682  
Qy 4385 CGGGAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG-CAG 4443  
Db 6683 CGGGAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGGAGCAG 6742  
Qy 4444 TTTTCTCCCTCGGACGGTTAGCCACTGAAAGAGGGGAAATACTTTTTCCTGCTGCAACTAT 4503  
Db 6743 TTTTCTCCCTCGGACGGCTAGCCACTGAAAGAGGGGAAATACTTTTTCCTGCTGCAACTAT 6802  
Qy 4504 CCAATGGAATTACTTAAACCCCTTCAACACCTTTTCACTTAGGCATCGATAGCACC 4563  
Db 6803 CCAATGGAATTACTTAAACCCCTTCAACACCTTTTCACTTAGGCATCGATAGCACC 6862  
Qy 4564 TCARATGGCCAAATCATTTATTTACTGGACACAGGCTTTTCAAACTATCAAGCARATAT 4623  
Db 6863 TCARATGGCCAAATCATTTATTTACTGGACACAGGCTTTTCAAACTATCAAGCARATAT 6922  
Qy 4624 CAGGCGCTGTGAATGTGCCARABAAATAATCCCTGCTTATCGCCAGCTCTTTCAGG 4683  
Db 6923 CAGGCGCTGTGAATGTGCCAGAGAAATAATCCCTGCTTATCGCCAGCTCTTTCAGG 6982  
Qy 4684 ARAACAAABAACAGGCCATTTACCTGARAARABACTGGCAACTGATTTTACCCACAGGCC 4743  
Db 6983 AGAACAAABAACAGGCCATTTACCTGARAARABACTGGCAACTGATTTTACCCACAGGCC 7042  
Qy 4744 AAACCTCAGGATTTTCACTAGTCTAGTCTGGGTARATCTTTTACCGGGTGGGCARAG 4803  
Db 7043 AAACCTCAGGATTTTCACTAGTCTGGGTARATCTTTTACCGGGTGGGCARAG 7102  
Qy 4804 GCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACATAGTTTCAATAATA 4863  
Db 7103 GCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACATAGTTTCAATAATA 7162  
Qy 4864 ATTCCAGATTCGAGACTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTTCAGGCC 4923  
Db 7163 ATTCCAGATTCGAGACTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTTCAGGCC 7222  
Qy 4924 ACAGTTAACCCAGGAGTATCCAGGCGTTAGTATACGATATCACTTACACTGCGCCTGA 4983  
Db 7223 ACAGTTAACCCAGGAGTATCCAGGCGTTAGTATACGATATCACTTACACTGCGCCTGA 7282  
Qy 4984 AGGCCACAGTCTCAGGGAGGTCGAGAAAATGAATAAAYACTCAAAAGGACATCTAAA 5043  
Db 7283 AGGCCACAGTCTCAGGGAGGTCGAGAAAATGAATAAAYACTCAAAAGGACATCTAAA 7342  
Qy 5044 AAGCAAAACCCAGGAAACCCACTCAGTGGCTGTCTGTGCTTATAGCCTTAAAAAGA 5103  
Db 7343 AAGCAAAACCCAGGAAACCCACTCAGTGGCTGTCTGTGCTTATAGCCTTAAAAAGA 7402  
Qy 5104 ATCTGCAACTTTCCCAAAAGGAGGCTTAGCCCATAGCAATGCTCTATGGAAGGCC 5163  
Db 7403 ATCTGCAACTTTCCCAAAAGGAGGCTTAGCCCATAGCAATGCTCTATGGAAGGCC 7462  
Qy 5164 TTCTAAACCAATGACTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGACAGATCACCT 5223  
Db 7463 TTCTAAACCAATGACTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGACAGATCACCT 7522  
Qy 5224 CTTTAGCCAAATATCAACAGTCTTTAAAAATTAACAAGGAACCTTATCCCTGAGAGAGG 5283  
Db 7523 CTTTAGCCAAATATCAACAGTCTTTAAAAATTAACAAGGAACCTTATCCCTGAGAGAGG 7582

QY 5284 GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTCYCTTAATTCCTC 5343  
DB 7583 GAAAGAACTATTCCACCCCTTGTGACATGGTATTAGTCAAGTCCCTCTTAATTCCTC 7642  
QY 5344 CATCCCTAGATACATCTCGGAGAGACCCCTACCCAGTCAATTTATYATACCCCACTGCGG 5403  
DB 7643 CATCCCTAGATACATCTCGGAGAGACCCCTACCCAGTCAATTTATCTACCCCACTGCGG 7702  
QY 5404 TTAAGTGGCTGGAGTGGAGTCTTGGATACATACATCTGAGTCAAAATCTCGGATATGC 5463  
DB 7703 TTAAGTGGCTGGAGTGGAGTCTTGGATACATACATCTGAGTCAAAATCTCGGATATGC 7762  
QY 5464 CAAAGGAACCTGAAATTCAGAGGAGACAACTAGTATCTCTGAACTCTAGAGATT 5523  
DB 7763 CAAAGGAACCTGAAATTCAGAGGAGACAACTAGTATCTCTGAACTCTAGAGATT 7822  
QY 5524 TGGCCCTGCTCTTCAAAACAACAGGAGGAAAGTAACATAAATCAATAATCCCCATG 5583  
DB 7823 TGGCCCTGCTCTTCAAAACAACAGGAGGAAAGTAACATAAATCAATAATCCCCATG 7881  
QY 5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCATGCA 5643  
DB 7882 GSCCTCCCTTATCATATTTTCTCTTACTTGTCTTTTACCTCTTCTCACTCTCACTGCA 7941  
QY 5644 CCCCTCCATGCGCTGATGACCACTAGTCTCCCTVACWAGAGTTTCTATGGAGATG 5703  
DB 7942 CCCCTCCATGCGCTGATGACCACTAGTCTCCCTTACCAAGAGTTTCTATGGAGATG 8001  
QY 5704 CAGCGTCCCGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCAACC 5763  
DB 8002 CAGCGTCCCGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCAACC 8061  
QY 5764 TTCACTGCCCAACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGCAT 8121  
DB 8062 TTCACTGCCCAACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGCAT 8121  
QY 5824 GCAAACTACTCAATTTGGACAGGAAATAATTAATCTAGTGTCTGAGAGACTTGA 5883  
DB 8122 GCAAACTACTCAATTTGGACAGGAAATAATTAATCTAGTGTCTGAGAGACTTGA 8181  
QY 5884 GTCACTGTCTGTGGACTTACTTCAACCAACTGATGCTGATGGGGGTGAGTTCAA 5943  
DB 8182 GTCACTGTCTGTGGACTTACTTCAACCAACTGATGCTGATGGGGGTGAGTTCAA 8241  
QY 5944 GATCAGCAAGAGAAACATGTAAAGAGTAAATCTCCAACTCAACCCGGGTACATGGC 6003  
DB 8242 GATCAGCAAGAGAAACATGTAAAGAGTAAATCTCCAACTCAACCCGGGTACATGGC 8301  
QY 6004 ACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCCAT 6063  
DB 8302 ACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCCAT 8361  
QY 6064 ACTCGCTGTAGGCTTATTAATACCACTCACTGGGCTCCATGAGTCTCGGCCAA 6123  
DB 8362 ACTCGCTGTAGGCTTATTAATACCACTCACTGGGCTCCATGAGTCTCGGCCAA 8421  
QY 6124 AACCTCTAACTGTGTGGATATGCTCCCTGAACTTCAAGCCATATGTTTCAATCCCT 6183  
DB 8422 AACCTCTAACTGTGTGGATATGCTCCCTGAACTTCAAGCCATATGTTTCAATCCCT 8481  
QY 6184 GTACCTGAAACATGGAAACATCTTACGACAGAAATAAACAACCTTCCGTTTGTAGGA 6243  
DB 8482 GTACCTGAAACATGGAAACATCTTACGACAGAAATAAACAACCTTCCGTTTGTAGGA 8541  
QY 6244 CCTCTGTGTTTCCAACTGGAAATAAACCATACCTCAACCTCACTGTGTAATAATTTAGC 6303  
DB 8542 CCTCTGTGTTTCCAACTGGAAATAAACCATACCTCAACCTCACTGTGTAATAATTTAGC 8601  
QY 6304 AATACTACATACAAACCACTCCCAATGATCAGGTGGTACTCTCTCCACACAAATA 6363  
DB 8602 AATACTACATACAAACCACTCCCAATGATCAGGTGGTACTCTCTCCACACAAATA 8661  
QY 6364 GTCTGCTACCTCAGGAATATTTTTTGTGTGTGTACCTCAGCCCTATCGTTGTTGAAT 6423

DB 8662 GTCTGCTACCTCAGGAATATTTTTTGTGTGTGTACCTCAGCCCTATCTGTTGAAT 8721  
QY 6424 GGCTCTTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCCYATGRCCTACTACACT 6483  
DB 8722 GGCTCTTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCCYATGRCCTACTACACT 8781  
QY 6484 GAAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCAATCTTT 6543  
DB 8782 GAAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCAATCTTT 8841  
QY 6544 CCTTTTGTATAGGAGCAGGAGTGTAGTGCACTAGTACTGGCAATGGCGGTATCACA 6603  
DB 8842 CCTTTTGTATAGGAGCAGGAGTGTAGTGCACTAGTACTGGCAATGGCGGTATCACA 8901  
QY 6604 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGTC 6663  
DB 8902 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGTC 8961  
QY 6664 GCGCACTCCCTGGTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTACTCTTCA 6723  
DB 8962 GCGCACTCCCTGGTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTACTCTTCA 9021  
QY 6724 AATCGAAGAGCTTGTAGACTTGTAACTGCTAAGAGTCAACTTAACTCCCTAGCAGCAGTACTCTTCA 6783  
DB 9022 AATCGAAGAGCTTGTAGACTTGTAACTGCTAAGAGTCAACTTAACTCCCTAGCAGCAGTACTCTTCA 9081  
QY 6784 GAAGATGCTGTATTATGTTAATCAATCCGAACTGCTCACTGAGAAAGTTTAAAGAAAT 6843  
DB 9082 GAAGATGCTGTATTATGTTAATCAATCCGAACTGCTCACTGAGAAAGTTTAAAGAAAT 9141  
QY 6844 CSAGATCGAATAACAAGTAKAGCAGAGACTTTCGAAACACTGGAACCTCGGGCCCTCCTC 6903  
DB 9142 CGAGATCGAATAACAAGTAKAGCAGAGACTTTCGAAACACTGGAACCTCGGGCCCTCCTC 9201  
QY 6904 AGGCRATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTTGCTA 6963  
DB 9202 AGCCAAATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTTGCTA 9261  
QY 6964 CTCTCTTTGGACCTGTATCTTTTAACTCTCTGTTAACTTTGCTCTCTCCAGAAATCGAA 7023  
DB 9262 CTCTCTTTGGACCTGTATCTTTTAACTCTCTGTTAACTTTGCTCTCTCCAGAAATCGAA 9321  
QY 7024 GCTGTAAACTACTAAATGGAGCCCAAGATGCACTCAAGACTAAGACTTACCCGACAGCC 7083  
DB 9322 GCTGTAAACTACTAAATGGAGCCCAAGATGCACTCAAGACTAAGACTTACCCGACAGCC 9381  
QY 7084 CTGAGCCGGCTGTATAGCCACGATCTGATGTTAATGACATCAAAAGGCAACCCCTCCTGAG 7143  
DB 9382 CTGAGCCGGCTGTATAGCCACGATCTGATGTTAATGACATCAAAAGGCAACCCCTCCTGAG 9441  
QY 7144 GAAATCTCAGCTGCACACCTCTACTACGCCCAATTCAGCAGGAGCAGTTAGAGCGT 7203  
DB 9442 GAAATCTCAGCTGCACACCTCTACTACGCCCAATTCAGCAGGAGCAGTTAGAGCGT 9501  
QY 7204 SGTGCGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 7263  
DB 9502 C-TGCGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 9560  
QY 7264 AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGTGAAC 7323  
DB 9561 AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCCTAGCTGGGAAGTGAAC 9620  
QY 7324 ACATCCACTTTAAACAAGGCTTGCACCTTAGTCTACACCTGACCAATCAGAGAGCTC 7383  
DB 9621 ACATCCACTTTAAACAAGGCTTGCACCTTAGTCTACACCTGACCAATCAGAGAGCTC 9680  
QY 7384 ACTAAAATGCTAATTTAGGCAAGAGCAGGAGTAAAGAAATGCAATCATYATTGCTMG 7443  
DB 9681 ACTAAAATGCTAATTTAGGCAAGAGCAGGAGTAAAGAAATGCAATCATYATTGCTMG 9740  
QY 7444 AGAGCAGCAGCAGGAGCAATGATCGGGATATAAACCAGTCTTCAGAGCCGCAACCG 7503

Db 9741 AGAGCACAGCAGGAGGCAATGATCGGGATATAAACCAAGTCTTCGAGCGGCAACGG 9800

Qy 7504 CAACCCCTTTGGGTCCCTCCCTTGGATGGAGCTCTGTTTCATGCTATTTCACCTCT 7563

Db 9801 CAACCCCTTTGGGTCCCTCCCTTGGATGGAGCTCTGTTTCATGCTATTTCACCTCT 9860

Qy 7564 ATTAATAATCTTGCACTGCR 7582

Db 9861 ATTAATAATCTTGCACTGCA 9879

RESULT 6

US-09-854-867-21

; Sequence 21, Application US/09854867

; Publication No. US20030224356A1

; GENERAL INFORMATION:

; APPLICANT: JOAN, KNOLL H

; APPLICANT: ROGAN, PETER K

; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING

; FILE REFERENCE: 30307

; CURRENT APPLICATION NUMBER: US/09/854,867

; CURRENT FILING DATE: 2003-05-08

; NUMBER OF SEQ ID NOS: 613

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 21

; LENGTH: 8523

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: repeat region

; LOCATION: (1)-(8523)

; OTHER INFORMATION: nerv17

US-09-854-867-21

Query Match 81.8%; Score 6124.6; DB 3; Length 8523;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

Qy 606 GATGGAAACGTTCCCGCAAGCAAAACGCGCTTAAGACGTATTCTGAGABAATTTGGGA 665

Db 1923 GATGGAAACGTTCCCGCAAGCAAAACGCGCTTAAGACGTATTCTGAGABAATTTGGGA 1982

Qy 666 MCAATTTGACCTCGACACTAAGAAAGAAAGACTTATTTCTCTGAGTCCGCGCTG 725

Db 1983 CCAATTTGACCTCGACACTAAGAAAGAAAGACTTATTTCTCTGAGTCCGCGCTG 2042

Qy 726 GCACCTCTGAGGAAGTATAAATTATAACACCATCTTACAGCTAGACACTTTTGTAGAA 785

Db 2043 GCACCTCTGAGGAAGTATAAATTATAACACCATCTTACAGCTAGACACTTTTGTAGAA 2102

Qy 786 AAGGCAAAATGGAGTGAAGTGCCATAGTAACAATCTTTCTTTTCAATAAGAGACAACCTCAC 845

Db 2103 AAGGCAAAATGGAGTGAAGTGCCATAGTAACAATCTTTCTTTTCAATAAGAGACAACCTCAC 2162

Qy 846 AATTATGTAAGTGTGTTTATGCGCTACAGGAGCCTTCAGAGTCTACCTCCCTATC 905

Db 2163 AATTATGTAAGTGTGTTTATGCGCTACAGGAGCCTTCAGAGTCTACCTCCCTATC 2222

Qy 906 CCAGCATCCCGACTCTTTCCCAATTAATAGGACCCCGCTTCAACCCCAATGGTCCAA 965

Db 2223 CCAGCATCCCGACTCTTTCCCAATTAATAGGACCCCGCTTCAACCCCAATGGTCCAA 2282

Qy 966 AAGGAGATAGACAAAGGGTAAACAGTGAAACCAAGAGTGGCAATATCCCAATATGA 1025

Db 2283 AAGGAGATAGACAAAGGGTAAACAGTGAAACCAAGAGTGGCAATATCCCAATATGA 2342

Qy 1026 CCCCTCCAGCAGTGGGAGGAGAAATTCGGCCAGCCAGTGCATGTCGCTTTT 1085

Db 2343 CCCCT-CCAAGCAGTGGGAGGAGAAATTCGGCCAGCCAGTGCATGTCGCTTTTTC 2401

Qy 1086 TCTCCAGACTTAAAGCAATAAAAAACAGACTTTAGGTAAATCTCAGATAAATCTGATGG 1145

Db 2402 TCTCCAGACTTAAAGCAATAAAAAACAGACTTTAGGTAAATCTCAGATAAATCTGATGG 2461

Qy 1146 CTATATTGRTGCTTTTCAAGGGTTAGGACAATTTCTTTGATCTGACATCGAGAGATATATA 1205

Db 2462 CTATATTGATGTTTTTCAAGGGTTAGGACAATTTCTTTGATCTGACATCGAGAGATATA-A 2520

Qy 1206 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCCACCAATACTGCAGCCT 1265

Db 2521 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCCACCAATACTGCAGCCT 2580

Qy 1266 GAGRGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGGATGACAAACAAGG 1325

Db 2581 GAGAGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATG--ATAGGATGACAACAG-AGG 2637

Qy 1326 AAAGANAATGATCCCAACAGCCAGCARGCAGTTCCAGCTTASACCTCATTTGGGGAC 1385

Db 2638 AAAGANAATGATCCCAACAGCCAGCARGCAGTTCCAGCTTASACCTCATTTGGGGAC 2694

Qy 1386 ACAGAAATCAGTAATCATGGAGATTTGGTCTGCAGACATTTGCTAACTTTGTGTCTASAA 1445

Db 2695 ACACAGATCAGACAT--GGAGATTTGGTCTGCAGACATTTGCTAACTTTGTGTCTAGNA 2753

Qy 1446 GGACTAAGGAAAACTTASGAGAAAACTTAYGAAATTACTCAATGATGTCCACATAACACA 1505

Db 2754 GGACTAAGGAAAACTTAGGAAG-AAAGTCTATGAATTTACTCAATGATGTCCACCATACACA 2812

Qy 1506 GGGGAAGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGGAGGCATTTGAGGAAGC 1565

Db 2813 -GGGAAGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGGAGGCATTTGAGGAAGC 2871

Qy 1566 GTGCTCTCTCTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCA 1625

Db 2872 GTGCTCTCTCTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCA 2931

Qy 1626 CTCAGTCAAGTGCAGACATTAAG-AAAAAATTTCAAAAGTCTGCCGTAGGCGCGAGCAAA 1684

Db 2932 CTCAGTCAAGTGCAGACATTAAGAAAAAATTTCAAAAGTCTGCCGTAGGCGCGAGCAAA 2991

Qy 1685 ACTTAGAAGCCTTATTTGAACCTTGGCACTTGGTCTTTTATAATAGAGATCAGGAGGAGC 1744

Db 2992 ACTTAGAAGCCTTATTTGAACCTTGGCACTTGGTCTTTTATAATAGAGATCAGGAGGAGC 3051

Qy 1745 AGGCGGAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTTAGTCAATGACCCCTCAGG 1804

Db 3052 AGGCGGAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTTAGTCAATGACCCCTCAGG 3111

Qy 1805 CAAGTGAATCTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAATAGG 1864

Db 3112 CAAGTGAATCTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAATAGG 3171

Qy 1865 GCTTGTCTTCCAGTGGGCTTCAAGGACACTTTTAAAAAAGATTTGTCGAAGTAGAAGTAG 1924

Db 3172 GCTTGTCTTCCAGTGGGCTTCAAGGACACTTTTAAAAAAGATTTGTCGAAGTAGAAGTAG 3231

Qy 1925 CGCGCCCTTCGTCAATGCGCTTATTTTCAAGGGAATCACTGGAAGGCCACCTGCGCCGAGG 1984

Db 3232 CGCGCCCTTCGTCAATGCGCTTATTTTCAAGGGAATCACTGGAAGGCCACCTGCGCCGAGG 3291

Qy 1985 GGACAAAAGGTCTTTTGTAGTCAAGGCCACTAACCCAGATGATCCAGCAGCAGACTGAGGG 2044

Db 3292 GGACAAAAGGTCTCTGTAGTCAAGGCCACTAACCCAGATGATCCAGCAGCAGACTGAGGG 3351

Qy 2045 TGCTTGGGGCAGCGCCATCCCATGCCATCCCTCACAGAGCCCTGGGTATGCTTGACC 2104

Db 3352 TGCTTGGGGCAGCGCCATCCCATGCCATCCCTCACAGAGCCCTGGGTATGCTTGACC 3411

Qy 2105 ATTGAGGGCCAGGAAGTTGTCTCTGGACACTGTGTGGGCTTCTTCTTAGTCTTACTCTTC 2164

Db 3412 ATTGAGGGCCAGG-AGGTTGTCTCTGGACACTGTGTGGGCTTCTTCTTAGTCTTACTCTTC 3470

Qy 2165 TGTCCCGGACAACTGTCTCTCCAGATCTGTCTACTATTCTTGAGGGGGTCCNTAAGACGGGCA 2224

Db 3471 TGTCCCGGACAACTGTCTCTCCAGATCTGTCTACTA-TCTGAGGGGGTCC-TAAGACGGGCA 3528



Db 5653 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTTCACGAGACAG 5712  
Qy 4444 TTTTCTCCCTCGGACGCTTAGCCACTGAAGAAGGAAATACTTTTGGCCCTGCAACTAT 4503  
Db 5713 TTTTCTCCCTCGGACGCTTAGCCACTGAAGAAGGAAATACTTTTGGCCCTGCAACTAT 5772  
Qy 4504 CCAATGGAAATTAATTAACACCTTTCATCAAAACCTTTCATCTAGGCATCGATAGCACCCA 4563  
Db 5773 CCAATGGAAATTAATTAACACCTTTCATCAAAACCTTTCATCTAGGCATCGATAGCACCCA 5832  
Qy 4564 TCARATGGCAATCATTAATTAATTAATTCAGGACAGGCTTTTCAAAATCAATCAAGCABATAT 4623  
Db 5833 TCAGATGGCAATCATTAATTAATTAATTCAGGACAGGCTTTTCAAAATCAATCAAGCABATAT 5892  
Qy 4624 CAGGSCCTGTGAATGTGCGACAGAAATTAATCCCTGCTCTATGCGCAAGCTCTTCAGG 4683  
Db 5893 CAGGSCCTGTGAATGTGCGACAGAAATTAATCCCTGCTCTATGCGCAAGCTCTTCAGG 5952  
Qy 4684 ARAACAAARAAACAGGCCATTACCTGGRARAARACTGGCAACTGATTTTACCCACAAGGCC 4743  
Db 5953 AGAACAAAGAACAGGCCATTACCTGGAGAAGACTGGCAACTGATTTTACCCACAAGGCC 6012  
Qy 4744 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATACATTTTCAAGGCTTGGGCAAG 4803  
Db 6013 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATACATTTTCAAGGCTTGGGCAAG 6072  
Qy 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCATAGTTTCATGAATA 4863  
Db 6073 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCATAGTTTCATGAATA 6132  
Qy 4864 ATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTCTTTCAGGCC 4923  
Db 6133 ATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTCTTTCAGGCC 6192  
Qy 4924 ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA 4983  
Db 6193 ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA 6252  
Qy 4984 AGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA 5043  
Db 6253 AGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA 6312  
Qy 5044 AAGCAAAACCCAGAAACCCACCTCACATGGCTGTCTGTGCTATAGCCTTAAAGAA 5103  
Db 6313 AAGCAAAACCCAGAAACCCACCTCACATGGCTGTCTGTGCTATAGCCTTAAAGAA 6372  
Qy 5104 ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAAGGCC 5163  
Db 6373 ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAAGGCC 6432  
Qy 5164 TTCAATACCAATGACCTTGTGCTTGACCCAGACGCAACTTAGTTGAGACATCACT 5223  
Db 6433 TTCAATACCAATGACCTTGTGCTTGACCCAGACGCAACTTAGTTGAGACATCACT 6492  
Qy 5224 CTTTAGCCAAATATCAACAAATTTCTTAAACATTAACAGGAACCTTCCCTCAGAGAGG 5283  
Db 6493 CTTTAGCCAAATATCAACAAATTTCTTAAACATTAACAGGAACCTTCCCTCAGAGAGG 6552  
Qy 5284 GAAAGAACTATTCCACCCWGTGACATGGTATFAGTCAAGTCCCTTCCTCTAATTC 5343  
Db 6553 GAAAGAACTATTCCACCCWGTGACATGGTATFAGTCAAGTCCCTTCCTCTAATTC 6612  
Qy 5344 CATCCCTAGATACATCCTGGGAAGGACCCCTACCCAGTCAATTTATACCCCACTGGG 5403  
Db 6613 CATCCCTAGATACATCCTGGGAAGGACCCCTACCCAGTCAATTTATACCCCACTGGG 6672  
Qy 5404 TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTGAAGTCAATCTCGATCTGC 5463  
Db 6673 TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTGAAGTCAATCTCGATCTGC 6732  
Qy 5464 CAAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTTCTGTGAACCTCTAGAGGATT 5523  
Db 6733 CAAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTTCTGTGAACCTCTAGAGGATT 6792

Qy 5524 TGGCCTGCTCTTCAACAAACACAGGAGGAAAGTAACTAAATCATAAATCCCCCATG 5583  
Db 6793 TGGCCTGCTCTTCAACAAACACAGGAGGAAAGTAACTAAATCATAAATCCCCCATG 6851  
Qy 5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCCTSTTTCACTCTCACTGCA 5643  
Db 6852 GSCCTCCCTTATCATATTTTCTCTKTACTGTTCTTTTACCCCTCTTTCACTCTCACTGCA 6911  
Qy 5644 CCCCCTCATGCGCTGTATGACACGAGTAGCTCCCTYACOMAGAGTTTCTATGGAGATG 5703  
Db 6912 CCCCCTCATGCGCTGTATGACACGAGTAGCTCCCTTACCAAGAGTTTCTATGGAGATG 6971  
Qy 5704 CAGCGTCCCGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCCCCACC 5763  
Db 6972 CAGCGTCCCGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCCCCACC 7031  
Qy 5764 TTCACTGCCACACCCCATATATGCCGCAACTGCTATCACTCTGCACTCTTTTGCATGAT 5823  
Db 7032 TTCACTGCCACACCCCATATATGCCGCAACTGCTATCACTCTGCACTCTTTTGCATGAT 7091  
Qy 5824 GCAATACTCATTTATGGACAGGAAATGATTAATCTTAGTCTCTGAGGACTTGA 5883  
Db 7092 GCAATACTCATTTATGGACAGGAAATGATTAATCTTAGTCTCTGAGGACTTGA 7151  
Qy 5884 GTCACTGCTCTTGGACTTACTTCAACCAAACTGGTATGCTGTAGGGGTGGAGTTCAA 5943  
Db 7152 GTCACTGCTCTTGGACTTACTTCAACCAAACTGGTATGCTGTAGGGGTGGAGTTCAA 7211  
Qy 5944 GATCAGGCAAGAGAAAAACATGTAAAGAACTAATCTCCCAACTCACCSGGGTACATGGC 6003  
Db 7212 GATCAGGCAAGAGAAAAACATGTAAAGAACTAATCTCCCAACTCACCSGGGTACATGGC 7271  
Qy 6004 ACCTTAGCCCTTACAAAGGACTAGTCTCTCAAAACTACATGAATAACCTTCGTTACCCAT 6063  
Db 7272 ACCTTAGCCCTTACAAAGGACTAGTCTCTCAAAACTACATGAATAACCTTCGTTACCCAT 7331  
Qy 6064 ACTGCCCTGTAAAGCTATTTAATACCACTCTACTGGGCTCCATGAGGTCTCGGCCCAA 6123  
Db 7332 ACTGCCCTGTAAAGCTATTTAATACCACTCTACTGGGCTCCATGAGGTCTCGGCCCAA 7391  
Qy 6124 AACCTACTAATCTGTGGATATGCTCCCTCCCTGAACTTCARGGCCATATGTTTCAATCCCT 6183  
Db 7392 AACCTACTAATCTGTGGATATGCTCCCTCCCTGAACTTCARGGCCATATGTTTCAATCCCT 7451  
Qy 6184 GTACTGAACTATGGAACAACTTCAGACAGAAATAAACAACACTTCGTTTGTAGTAGGA 6243  
Db 7452 GTACTGAACTATGGAACAACTTCAGACAGAAATAAACAACACTTCGTTTGTAGTAGGA 7511  
Qy 6244 CCTCTGTTTCCCAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAATAATTTAGC 6303  
Db 7512 CCTCTGTTTCCCAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAATAATTTAGC 7571  
Qy 6304 AATACTACATACAAACCAACTCCCAATGCACTCAGGTGGGTAACTCTCCCAACACAAATA 6363  
Db 7572 AATACTACATACAAACCAACTCCCAATGCACTCAGGTGGGTAACTCTCCCAACACAAATA 7631  
Qy 6364 GTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCTATGCTGTTTGTGAAT 6423  
Db 7632 GTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCTATGCTGTTTGTGAAT 7691  
Qy 6424 GGCTCTTCAGGAATCTATGCTTCTCTCATTTCTTAGTGCCCTTATGACCACTACACT 6483  
Db 7692 GGCTCTTCAGGAATCTATGCTTCTCTCATTTCTTAGTGCCCTTATGACCACTACACT 7751  
Qy 6484 GAAAGAAATTTATACAGTTATGCTATATCTAAGCCCGCAACAAAGAGTACCCATTTCT 6543  
Db 7752 GAAAGAAATTTATACAGTTATGCTATATCTAAGCCCGCAACAAAGAGTACCCATTTCT 7811  
Qy 6544 CTTTTTGTATAGGACGAGGAGTGTAGGTGCACTAGGTGCACTAGGTGCACTTGGCGGTATCACA 6603  
Db 7812 CTTTTTGTATAGGACGAGGAGTGTAGGTGCACTAGGTGCACTTGGCGGTATCACA 7871







1506 GGGGAAGGGAAGAAATCCTACTGCTTTCTGGAGAGACTAAGGAGGCAATTGAGGAAGC 1565  
Db |||||  
2813 -GGGAAGGGAAGAAATCCTACTGCTTTCTGGAGAGACTAAGGAGGCAATTGAGGAAGC 2871  
Qy |||||  
1566 GTGCTCTCTGTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCA 1625  
Db |||||  
2872 GTGCTCTCTGTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCA 2931  
Qy |||||  
1626 CTCAGTCAGCTCAGACATAG-AAAAAATCTCAAAAGTCTGCGTAGGCGCGGAGCAAA 1684  
Db |||||  
2932 CTCAGTCAGCTCAGACATAGAAAAAATCTTCAAAAGTCTGCGTAGGCGCGGAGCAAA 2991  
Qy |||||  
1685 ACTTAGAAACCCCTATTGAACCTTGGCAACVTCGGTTTTTATATAATAGAGATCAGAGGAGC 1744  
Db |||||  
2992 ACTTAGAAACCCCTATTGAACCTTGGCAACCTCGGTTTTTATATAATAGAGATCAGAGGAGC 3051  
Qy |||||  
1745 AGGCGGAACAGGACAAAACGGGATTAAAAAAGGCGCACCGCTTTAGTCATGACCCCTCAGG 1804  
Db |||||  
3052 AGGCGGAACAGGACAAAACGGGATTAAAAAAGGCGCACCGCTTTAGTCATGACCCCTCAGG 3111  
Qy |||||  
1805 CAACTGAGCTTTGGAGGCTCTCGAAAGGGAAGCTGGGCAAAATGGAATGCTTAATAGG 1864  
Db |||||  
3112 CAACTGAGCTTTGGAGGCTCTCGAAAGGGAAGCTGGGCAAAATGGAATGCTTAATAGG 3171  
Qy |||||  
1865 GCTTGCTTCCAGTGGGCTCTACAGGACACTTTTAAAAAGATTGTCCAAAGTAGAAGTAAG 1924  
Db |||||  
3172 GCTTGCTTCCAGTGGGCTCTACAGGACACTTTTAAAAAGATTGTCCAAAGTAGAAGTAAG 3231  
Qy |||||  
1925 CCGCCCCCTTGGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 1984  
Db |||||  
3232 CCGCCCCCTTGGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 3291  
Qy |||||  
1985 GGACAAAGGCTTTTGAAGTCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG 2044  
Db |||||  
3292 GGACAAAGGCTTCTGAGTCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG 3351  
Qy |||||  
2045 TGCTGGGCGACGCGCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACC 2104  
Db |||||  
3352 TGCTGGGCGACGCGCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACC 3411  
Qy |||||  
2105 ATTGAGGGCGAGGAAGTTGTCTCTCGGACACTGTGTGCGGTCTTCTTAGTCTTACTCTTC 2164  
Db |||||  
3412 ATTGAGGGCGAGG-AGGTTGTCTCCTGGACACTGTGTGCGGTCTTCTTAGTCTTACTCTTC 3470  
Qy |||||  
2165 TGTCCCGGACAACTGTCTCCAGATCTGTCACTAATTCTAGGGGGTCCNTAAGACGGGCA 2224  
Db |||||  
3471 TGTCCCGGACAACTGTCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA 3528  
Qy |||||  
2225 GTCACTAGATACTTTTCCAGCCACTAAGTTATGAACCTGGGGAGCTTTATTTCTTTTAC 2284  
Db |||||  
3529 GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTGGGGAGGCTTTATTTCTTTTAC 3586  
Qy |||||  
2285 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTGTAGGAGAGACATTTCTAGCAA 2344  
Db |||||  
3587 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTGTAGGAGAGACATTTCTAGCAA 3646  
Qy |||||  
2345 AAGCAGGGCCATTATACACCTGAACATAGGAAGGAACAACCGTTGTGTGTCGCTG 2404  
Db |||||  
3647 AAGCAGGGCCATTATACACCTGAACATAGGAAGGAACAACCGTTGTGTGTCGCTG-CCGCTG 3705  
Qy |||||  
2405 CTTGAGGAAGGAATTAATCCTGAGTCTGGGCAACAGAGGACATATGAGCGGCCAAA 2464  
Db |||||  
3706 CTTGAGGAAGGAATTAATCCTGAGTCTGGGCAACAGAGGACATATGAGCGAG-CAAA 3764  
Qy |||||  
2465 GAATGCCGCTGCTGTTTCAAGTTTAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA 2524  
Db |||||  
3765 GAATGCCGCTGCTGTTTCAAGTTTAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA 3824  
Qy |||||  
2525 GTACCCCTCTCAGACCCAGGCCCAACAAGATTTCAAAAGATTGTTAAAGCACTTAAAGC 2584  
Db |||||  
3825 GTACCCCTCTCAGACCCAGGCCCAACAAGACTTCCAAAAGATTGTTAAAGCACTTAAAGC 3884  
Qy |||||  
2585 CCAAGGCTTAGTAAAAACCATGCAATACTCCCTGCAGTAATCCGTAGTGGATTGAGGAGG 2644

3885 CCAAGGCTTAGTAAAAACCATGCAGTAAACCCCTGCAGTACTCC-----AATTTTAGGAG 3937  
Qy |||||  
2645 CACAGAAACCCAGTGGGACAGTGGAGGTTAGTCAAGATCTCAGGATTTATCAATGAGGCG 2704  
Db |||||  
3938 TACAGAAACCCAAACAGACAGTGGGA-GGTTAGTCAAGATCTCAGGATTTATCAAT-GAGGC 3995  
Qy |||||  
2705 GGTGTGCTTTTATACCCAGCTGTACCTAGCCCTTATCTGMYTTTCCCAAAATACCAGA 2764  
Db |||||  
3996 TGTGTGCTTCTATAGCCAGCTGTACCTAGCCCTTATCTGCTTTTCCCAATACCAGA 4055  
Qy |||||  
2765 GGAACGACAGTGGTTTACASTCTCTGGACCTTTMAGGATGCTTTCTTCTGCATCCCTGTACA 2824  
Db |||||  
4056 GGAACGACAGTGGTTTACAGTCTCTGGACCTTCAGGATGCTTTCTTCTGCATCCCTGTACA 4115  
Qy |||||  
2825 TCCTGACTCTCAATCTTCTTGTGCTTTGAAAGATACCTTCAAAACCCARCATCTCAACTCAC 2884  
Db |||||  
4116 TCCTGACTCTCAATCTTCTTGTGCTTTTGAAGATACCTTCAAAACCCCACTCAACTCAC 4175  
Qy |||||  
2885 CTGACTTCTTTTACCCCAAGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAATTAGC 2944  
Db |||||  
4176 CTGACTTATTTTACCCCAAGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAATTAGC 4235  
Qy |||||  
2945 CCAAGACTTGAGYCARVTMTCATACCTGGGACACTCTTGTCTTCTGRTAGTGGATGATTT 3004  
Db |||||  
4236 CCAAGACTTGAGCCAACTCTCATCTGGACA--CTTGTCTTGGTAGTGGATGATTT 4293  
Qy |||||  
3005 ACTTTTTCGCTTTCAGAAACCTTGTGCAATCAAGCCACCCAGGCTCTTMAATTT 3064  
Db |||||  
4294 ACTTTTTCGCTTTCAGAAACCTTGTGCAATCAAGCCACCCAGGCTCTTCAATTT 4353  
Qy |||||  
3065 CCTGCTACCTGCTGCTACAGGTTTCCAAACCCARAGCTCARTCTCTCTCACAGCAGT 3124  
Db |||||  
4354 CCTGCTACCTGCTGCTACAGGTTTCCAAACCCARAGGCTCACTCTCTCACAGCAGT 4413  
Qy |||||  
3125 TAAATACCTTAGGCTCTAATATTTATCCAAAGGCAACARAGGCTCTCAGTAGGAAYRATCCA 3184  
Db |||||  
4414 ----TACTTAGGCTTAAATTTATCCAAAGGCAACARAGGCTCTCAGTAGGAACACATCCA 4469  
Qy |||||  
3185 GCCTATACCTGCTTATCTCTATCTCAAAACCTTAAAGCAACTAAGRRRTTCTCTGGGRT 3244  
Db |||||  
4470 GCCTATACCTGCTTATCTCTATCTCAAAACCTTAAAGCAACTAAGGGGATCTCTTGGGCT 4529  
Qy |||||  
3245 AAYAGGTTCTCCGAAWATGATTTCCCGAGTWTGCAAAATAGCCAGGCTCATTAWATA 3304  
Db |||||  
4530 AATAGGTTCTCCGAAWATGATTTCCCGAGTWTGCAAAATAGCCAGGCTCATTAWATA 4588  
Qy |||||  
3305 CASTAATTAAGGAACTCAGAAAGCCAAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364  
Db |||||  
4589 CACTAATTAAGGAACTCAGAAAGCCAAATACCCATTTAGTAGATGGACAACTGAAGTAG 4648  
Qy |||||  
3365 AAGTGGCTTTCAGGCCCCCTTAAAGAGGCTTAAACCCAGYCCAGTGTTAAGYTTGCC 3424  
Db |||||  
4649 AAGTGGCTTTCAGGCCCC-----TAAACCCAGCCCCAGTGTTAAGYTTGCC 4694  
Qy |||||  
3425 AACGGGCAAGACTTTTSTTATATRTCAAGAAACCAAGAAAGCTCTGAGGCTCC 3484  
Db |||||  
4695 AACGGGCAAGACTTTTCTTATATGTACAG-AAAAAACAGGAATAGCTCTGAGGCTCC 4753  
Qy |||||  
3485 TTACACAGTCCRAGGGAYGAGCTTGCAACCVRTGGRYACCTGASTAAGGAAAYTGATG 3544  
Db |||||  
4754 TTACACAGTCCGAGGGATGAGCTTGCAACCTGTGGCATACTGACTAAGGAAATGATG 4813  
Qy |||||  
3545 TAGTGGCAAAAGGTTGRCYCTCATTTTAYGGGTAGTGGTGCACTAGCAGYKTAGTAT 3604  
Db |||||  
4814 TAGTGGCAAAAGGTTGACCTCATTTTACGGGTAGTGGTGCACTAGCAGTCTTAGTAT 4873  
Qy |||||  
3605 CTGAGCAGTTAAATATACAGGAGAGATCTTACTGTGGGACATCTCATGAGTCA 3664  
Db |||||  
4874 CTGAGCAGTTAAATATACAGGAGAGATCTTACTGTGGGACATCTCATGAGTCA 4933  
Qy |||||  
3665 AYRGCACTACTCACTCTAAAGGAGACTTGTGGCTCTCAGACAACYGTTTACTTAAATRTC 3724  
Db |||||

4934	DB	ATGGCATACTCAGTCTGCTAAAGGAGACTTGTGGCTGTGCAGACAACTGTTTACTTTAAATGTCTC	4999
3725	QY	AGGCTCTATTACTTTGAARGGCCAGTGTGCTGCRACCTGTGCACCTTGTGCAACTCTTTAAACCCAG	3784
4994	DB	AGGCTCTANTACTTTGAAGGGCCAGTGTGGGACTGTGCACCTTGTGCAACTCTTTAAACCCAG	5053
3785	QY	YCNCATTTCTTCCAGACAATGAAGAAAGATARAAYATTAATCTGTCAACAARTAAATTTCTC	3844
5054	DB	CCAATTTCTTCCAGACAATGAAGAAAGATATAAACATAACTGTGTCAACAAGTAATTTCTC	5113
3845	QY	AAACCTATGCCACTCGAGGGGACCCTTGTAGAGTTTCCVTTTGACTGATCCYGCACCTTCAAC	3904
5114	DB	AAACCTATGCCACTCGAGGGGACCCTTTTAGAGTTTCTTTTGACTGATCCCGACC-TCCAC	5172
3905	QY	TTGTATACTGATGAAAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGTGGGGTATGCAATG	3964
5173	DB	TTGTATACTGATGAAAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGTGGGGTATGCAATG	5232
3965	QY	GTCACTGATTAATGCAATYTTGAAGTAATCCCTCTACTCCAGGAACCTAGTGTCTYAGCTR	4024
5233	DB	GTCACTGATTAATGCAATYTTGAAGTAATCCCTCTACTCCAGGAACCTAGTGTCTYAGCTR	5292
4025	QY	GCAGAACTAATAGCCYTCVTTKGGCCACTAGAAATTTAGGAGAGGAAAAAGGGYAAATATA	4084
5293	DB	GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTTAGGAGAGGAAAAAGGGCAATATA	5352
4085	QY	TATACAGACTCTRAATATGCTYACCTAGTCTNTCCATGCCCATGMRGCAATATATGSARAGAA	4144
5353	DB	TATACAGACTCTRAATATGCTTACTAGTCTCTCCATGCCCATGCGACCAATATGGAAGAA	5412
4145	QY	AGGAAATTCCTTAACCTTCYAGRGACACCTTATCMACATCAGGAGCCATTTAGGABATTA	4204
5413	DB	AGGAAATTCCTTAACCTTCGAGAGAACCTTATCAAACTCAGGAAGCCATTTAGGAAATTA	5472
4205	QY	TTATYTGCGWGTACAGAAACCTTARAGAGTGTGAGTCTTTACACTGCTGGGGTCTATCANAA	4264
5473	DB	TTATYTGCGTGTACAGAAACCTTAAGAGGTGGGAGTCTTTACACTGCGGGGTCTATCAGAA	5532
4265	QY	GGAAAGRAAGGGAAATASAGRGAAATYTGCAAGCAKATATTCGAAGCMAAAGAGCTGCA	4324
5533	DB	GGAAAGRAAGGGAAATAGAAGAGAACTGCCAAGCAGATATTCGAAGCAAGAGCTGCA	5592
4325	QY	AGGCAGAACCTTCAATTAGAAATGCTTTAAACTCTTCTTTAGTATAGGGTAATCCCTTC	4384
5593	DB	AGGCAGAACCTTCAATTAGAAATGCTTTATAAAACAACCCCTAGTATAGGGTAATCCCTTC	5652
4385	QY	CGGGAAACCAAGCCCCAGTACTCAGCAGGAGAACAGAAATGGGGAACCTTCACAGG-CAG	4443
5653	DB	CGGGAAACCAAGCCCCAGTACTCAGCAGGAGAACAGAAATGGGGAACCTTCACAGGACAG	5712
4444	QY	TTTTTCTCCCTTCGGGACGGTTAGCCACTGAAGAGGGGAAATATCTTTTGCTCTCAACTAT	4503
5713	DB	TTTTTCTCCCTTCGGGACGGTTAGCCACTGAAGAGGGGAAATATCTTTTGCTCTCAACTAT	5772
4504	QY	CCAATGGAAATATCTTAAACACCTTCCATCAAACTTTTCACTTAGGCATCGATAGCACCCA	4563
5773	DB	CCAATGGAAATATCTTAAACACCTTCCATCAAACTTTTCACTTAGGCATCGATAGCACCCA	5832
4564	QY	TCARATGGCCAAATCATTTTACTTGAACAGGCTTTTCAAACTATCAGCARATAT	4623
5833	DB	TCARATGGCCAAATCATTTTACTTGAACAGGCTTTTCAAACTATCAGCARATAT	5892
4624	QY	CAGGGCTGTGAATGTGCCARAAATAATCCCTGCGCTYATCGCCAGAGCTCTTTCAGG	4683
5893	DB	CAGGGCTGTGAATGTGCCARAAATAATCCCTGCGCTTATCGCCAGAGCTCTTTCAGG	5952
4684	QY	ARAAACAARAACAGGCCATTACCTGRRARAARACTTGGCACTGATTTTATCCCAAGCCCC	4743
5953	DB	ARAAACAARAACAGGCCATTACCTGRRARAARACTTGGCACTGATTTTATCCCAAGCCCC	6012
4744	QY	AAACCTCAGGGATTTCAATATCTATAGTCTGGGTATATCTTTTCAAGGGTGGGCAAG	4803
6013	DB	AAACCTCAGGGATTTCAATATCTATAGTCTGGGTATATCTTTTCAAGGGTGGGCAAG	6072

QY	4804	GC	TTTCCCTCTG	TACG	ACAG	AAAGGCC	CCAA	GAGGTAA	TAAAGGC	AC	TAGTTC	ATGAA	TA	4863					
DB	6073	GC	TTTCCCTCTG	TAGG	ACAG	AAAGGCC	CCAA	GAGGTAA	TAAAGGC	AC	TAGTTC	ATGAA	TA	6132					
QY	4864	AT	TCCAGATT	CGGAC	TTTCC	CCGAGG	CTTAC	CAGAGT	GACA	TATAG	CCCTG	TTTCC	AGGCC	4923					
DB	6133	AT	TCCAGATT	CGGAC	TTTCC	CCGAGG	CTTAC	CAGAGT	GACA	TATAG	CCCTG	TTTCC	AGGCC	6192					
QY	4924	AC	AGTAA	CC	CAGG	GAGTAT	CC	CAGG	CGTT	TAGG	TATAC	GA	TATAC	TA	4983				
DB	6193	AC	AGTAA	CC	CAGG	GAGTAT	CC	CAGG	CGTT	TAGG	TATAC	GA	TATAC	TA	6252				
QY	4984	AG	GC	AC	AGT	CCTC	CAG	GG	AGGT	CG	AGAA	TG	AA	TA	5043				
DB	6253	AG	GC	AC	AGT	CCTC	CAG	GG	AGGT	CG	AGAA	TG	AA	TA	6312				
QY	5044	A	G	C	A	A	A	A	A	A	A	A	A	A	5103				
DB	6313	A	G	C	A	A	A	A	A	A	A	A	A	A	6372				
QY	5104	AT	CTG	CA	ACT	TTTCC	CC	AA	AG	CAG	GA	CTTTAG	CC	CA	TA	5163			
DB	6373	AT	CTG	CA	ACT	TTTCC	CC	AA	AG	CAG	GA	CTTTAG	CC	CA	TA	6432			
QY	5164	TT	C	A	T	A	C	A	A	A	A	A	A	A	A	5223			
DB	6433	TT	C	A	T	A	C	A	A	A	A	A	A	A	A	6492			
QY	5224	CT	TTAG	CC	AA	TAT	CA	A	CA	AGT	TTCT	T	A	A	A	5283			
DB	6493	CT	TTAG	CC	AA	TAT	CA	A	CA	AGT	TTCT	T	A	A	A	6552			
QY	5284	G	A	A	A	G	A	A	C	T	T	A	A	A	C	TA	5343		
DB	6553	G	A	A	A	G	A	A	C	T	T	A	A	A	C	TA	6612		
QY	5344	C	A	T	C	C	T	A	G	A	T	A	C	C	C	T	TA	5403	
DB	6613	C	A	T	C	C	T	A	G	A	T	A	C	C	C	T	TA	6672	
QY	5404	T	T	A	A	G	T	G	G	A	G	T	G	G	A	G	TA	5463	
DB	6673	T	T	A	A	G	T	G	G	A	G	T	G	G	A	G	TA	6732	
QY	5464	C	A	A	A	G	A	A	C	T	G	A	A	A	T	C	C	TA	5523
DB	6733	C	A	A	A	G	A	A	C	T	G	A	A	A	T	C	C	TA	6792
QY	5524	T	G	C	C	T	G	C	T	C	T	T	C	A	A	A	C	TA	5583
DB	6793	T	G	C	C	T	G	C	T	C	T	T	C	A	A	A	C	TA	6851
QY	5584	G	S	C	C	T	C	T	T	A	T	T	T	T	T	T	T	TA	5643
DB	6852	G	S	C	C	T	C	T	T	A	T	T	T	T	T	T	T	TA	6911
QY	5644	C	C	C	C	T	C	A	G	C	T	G	T	A	C	C	C	TA	5703
DB	6912	C	C	C	C	T	C	A	G	C	T	G	T	A	C	C	C	TA	6971
QY	5704	C	A	G	C	G	T	A	T	T	G	A	T	G	A	T	G	TA	5763
DB	6972	C	A	G	C	G	T	A	T	T	G	A	T	G	A	T	G	TA	7031
QY	5764	T	T	C	A	C	T	G	C	C	C	A	A	C	T	G	C	TA	5823
DB	7032	T	T	C	A	C	T	G	C	C	C	A	A	C	T	G	C	TA	7091
QY	5824	G	C	A	A	A	T	A	T	T	G	G	A	A	A	T	A	TA	5883
DB	7092	G	C	A	A	A	T	A	T	T	G	G	A	A	A	T	A	TA	7151



224 GATCTCTGAGGGTGTCCGTGTCTCTCTGATCCAGCGARGCGCCCATTTGCCGTCCCAA 283  
2199 GATCCGCGAGGGTGTCTGTGATCTCTGATCCAGCGAGCGGCCCATTTGCCAGCCCAA 2258  
284 TTGGGCTAAAGGCTTGCCATGTGNCCTGCAAGCGCTAAGTGCCTGGGTTTGTCTTAATTGA 343  
2259 TCGGGCTAAAGGCTTGCCATTTGCTCTGCAAGGTTAAGTGCCTGGGTTGGTCTGTAATCGA 2318  
344 GCTGACACTANTACCTGGGTTCCATGGTTCTCTCTGTCACCCAGCGCTTCTAATA-KA 402  
2319 GCTGACACTTGTCACTGGGTTCCAGGTTCTCTTCCGTGACCCCATGGCTTCTTAATAGGA 2378  
403 ACTAATAACCTTACCACAT-GGGCCAAAGATTCCATCTCTT-GGAATCCGTGAGSSCAACG 460  
2379 ACTATTAACTTACCACATGGGCCCAAGATTCATCTCTTGGAAATCCGTGAGGCCAAG 2438  
461 AACTCCAGGTCAGAGATACAGRGCTTGGCCACCATCTTGAAGCGGCTCTCTACCRCTTT 520  
2439 AACCCAGGTCAGAGATACAAGGCTTGGCCACCATCTTGAAGCGGCCCTGCGCCATCTT 2498  
521 GGAAGTGGTTACCAACCATCTTGGGAGCTCTGTGAGCAAGGACCCCGCGTTRACATTTTG 580  
2499 GGAAGCGGCTTGGCCACCATCTTGGGAGCTCTGTGCACAAGGACCCCGCGTAAACATTTTG 2558  
581 GCRACCAMSRACGACATCCMAAGTGTGAGGAACGTTCCCGCAAGACAAAAACGCCCC 640  
2559 GCACCAAGCAAGGACATCCAAAGTGTGAGGAACGTTCCCGCAAGACAAAAACGCCCC 2618  
641 TAAGAGCTATTCTGGARAATTGGGACAAATTTGACCCCTCAGACACTAA-GAAAAGAAACGA 699  
2619 TAAGAGCTATTCTGGAGAAATTTGGACCAATTTGACCCCTCAGACACTTAAGAAAGAAACGA 2678  
700 CTTATATTCTTCTGAGTGCCGCTGGCACTCTGAGGGAAGTATAAATTTATAACACCAT 759  
2679 CTTATATTCTTCTGAGTGCCGCTGGCACTCTGAGGGAAGTATAAATTTATAACACCAT 2738  
760 CTTACAGCTAGACVCTCTTTGTAGAA--AAGGCAATGGAGTGAAGTGCCTAAGTACA 816  
2739 CTTACAGCTAGACTCTTTTGTAGAAAGAAAGTAAATGGAGTGAAGTGTATATGTACA 2798  
817 AACTTTCTTTTCAATTAAGAGACAACTCACAAATTA-TGTAAAAAGTGTGATTTATGCC-- 873  
2799 AACTTTCTTTTCAATTAAGAGACAACTCACAAATTTATTTGTAAAAAGTGTGATTTATGCCCT 2858  
874 TACAGGAAGCTTTCAGAGTCTACCTCCCTATCCC-----AGCATCCCAGTCTCTTCCCCA 929  
2859 ACAGGAAGGCTTTCAGAGTCTACCTCCCTATCCCAGCATCCCTCCCACTCTTCTCCA 2918  
930 MYTAATAAGACCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAAC 989  
2919 ACTATTAAAGGACCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAAC 2978  
990 AGTGAACCAAGAGTGCCAAATATTTCCCAATATGACCCCTCCCAAGCATGTGGAGGAAG 1049  
2979 AGTGAACCAAGAGTGCCAAATATTTCCCAATATGACCCCTCCCAAGCATGTGGAGGAAG 3038  
1050 AGAATTCGGGCCACGACAGTGTGCTGCTTTTCTCCAGACTTAAAGCAATAAA 1109  
3039 AGAATTCGGGCCACGACAGTGTGCTGCTTTTCTCTCCAGACTTAAAGCAATAAA 3098  
1110 AACAGACTTAGGTAATTTCTCAGATAAYCCTGATGCTATATTTGRTGTTTTTACAAGGGTT 1169  
3099 AACAGACTTAGGTAATTTCTCAGATAAACCTGATGCTATATTTGRTGTTTTTACAAGGGTT 3158  
1170 AGGACAAATTTTGTATCTGACATGGAGAGATATATATGTCATGCTGCTTAATTCAGACATAA 1229  
3159 AGGACAAATTTTGTATCTGACATGGAGAGATATATATGTCATGCTGCTTAATTCAGACATAA 3218  
1230 CCCCAATGAGAGAAGTCCACCATAACTGACGCTGAGGTTTGGCGATCTCTGCTATC 1289  
3219 CCCCAATGAGAGAAGTCCACCATAACTGACGCTGAGGTTTGGCGATCTCTGCTATC 3278  
1290 TCAGTCAGTCAATGGATANGATGACAAAGAGAAAGANAATGATTTCCCAAGGCC 1349

3279 TCAGTCAGTCAATGGATA--GGATGACACAG-AGAAAAGAGACGATTTCCCAAGGCC 3336  
1350 AGCARGCAGTTCAGTCTASACCTCATTTGGGGACACAGAAATCAGTAAACATGGGAGAT 1409  
3337 AGCARGCAGTTCAGTCTAGACCCCTCATTTGGGGACACAGAAATCAGTAAACATGGGAGAT 3396  
1410 TGTGCTGACAGACATTTGCTAACTTGTGCTCTASAGGACTAAGGAAAACTASGAGAAAA 1469  
3397 TGTGCTGACAGACATTTGCTAACTTGTGCTCTCAAGGACTAAGGAAAACTACGAGAAAA 3456  
1470 RTCTATGAATTACTCAATGATGTCCACCAATAACACAGGGGAGGAGAAAACTCTACTG 1529  
3457 ATCTAGAAATTACTCAATGATGTCCACCAATAACACAGGGGAGGAGAAAACTCTACTG 3516  
1530 CTTTCTTGGAGACATAAGGGAGGCAATTGAGGAAGGCTGCTCTCTGTCACCTGACTCTT 1589  
3517 CTTTCTTGGAGACATAAGGGAGGCAATTGAGGAAGGCTGCTCTCTGTCACCTGACTCTT 3576  
1590 CTGAAGGCCAACTAATCTTTAAAGCGTAAGTTTATCACTCAGTCTGACGATGACATTAAGAA 1649  
3577 CTGAAGGCCAACTAATCTTTAAAGCGTAAGTTTATCACTCAGTCTGACGATGACATTAAGAA 3636  
1650 AAAAATTTCAAAAAGTCTGCCGTAGGCCCGGAGCAAACTTAGAAAACTTGAATTTGGC 1709  
3637 AAAAATTTCAAAAAGTCTGCCGTAGGCCCGGAGCAAACTTAGAAAACTTGAATTTGGC 3696  
1710 AACCTCGGTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAAACAGACAAAACTGGGATTA 1769  
3697 AACCTCGGTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAAACAGACAAAACTGGGATTA 3756  
1770 AAAAAAGGCCAACCGCTTTAGTCATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTGAAA 1829  
3757 AAAAAAGGCCAACCGCTTTAGTCATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTGAAA 3816  
1830 AAGGGAAGGCTGGGCAAAATTTGAATGCTTAATAGGCTTGTCTCAGTGGGCTCTACAAG 1889  
3817 AAGGGAAGGCTGGGCAAAATTTGAATGCTTAATAGGCTTGTCTCAGTGGGCTCTACAAG 3876  
1890 GACACTTTAAAAAGATTTGTTCAAAGTAGAAGTAAAGCGCCCTTCTGTCATGCGCCCTTAT 1949  
3877 GACACTTTAAAAAGATTTGTTCAAAGTAGAAGTAAAGCGCCCTTCTGTCATGCGCCCTTAT 3936  
1950 TTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGCTCTTTTGAAGTCTAGAAG 2009  
3937 TTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGCTCTTTTGAAGTCTAGAAG 3996  
2010 CCACTAACAGATGATCCAGCAGGAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATG 2069  
3997 CCACTAACAGATGATCCAGCAGGAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATG 4056  
2070 CCATCAACCTCAGAGCCCTGGGTATGCTTGACCAATGAGGGCCAGGAGGTTGTCTCC 2129  
4057 CCATCAACCTCAGAGCCCTGGGTATGCTTGACCAATGAGGGCCAGGAGGTTGTCTCC 4116  
2130 TGGACACTGGTGGGCTTTCTTAGTCTTACTCTTCTGTCGCCGACAACTGTCTCTCAGAT 2189  
4117 TGGACACTGGTGGGCTTTCTTAGTCTTACTCTTCTGTCGCCGACAACTGTCTCTCAGAT 4176  
2190 CTGTCACTATTCTGAGGGGCTCCNTAAGACGGGCACTCATAGATACTTTTCTCCAGGCCA 2249  
4177 CTGTCACTATTCTGAGGGGCTCC-TAAGACGGGCACTCATAGATACTTTTCTCCAGGCCA 4235  
2250 CTAAGTATGAAGTGGGAGCTTTTATTCTTTTCACTGCTTTCTTAATTTATGCTGAAAG 2309  
4236 CTAAGTATGAAGTGGGAGCTTTTATTCTTTTCACTGCTTTTCTAAATTTATGCTGAAAG 4295  
2310 CCCCACTACTGTTTGGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTTATACACCTGAA 2369  
4296 CCCCACTACTGTTTGGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTTATACACCTGAA 4355  
2370 CATAGGAGAGGAAACCCGCTTTGTTGTCCTGCTTGGAGGAAGGAATTAATCTCTGAAG 2429

Db 4356 CATAGGAGAGGAAACACCGTTTGTGT - CCCTGCTTGAGGAGGAATTAATCCTGAAG 4414  
QY  
2430 TCTGGGACAGAGCAATATGACGACGACCAAGAGATGCCGCTCCCTGTTCAAGTTAAA 2489  
Db  
4415 TCTGGGCAACAGAGGCAATATGACGACGACCAAGAGATGCCGCTCCCTGTTCAAGTTAAA 4474  
QY  
2490 CTAAAGGATTCACACTTCCCTTCCCTACCAAGGCAAGTACCCCTCAGACCCCAAGGCCCAA 2549  
Db  
4475 CTAAAGGATTCGCTTCCCTTCCCTACCAAGGCAATACCCCTTAGACCTGAGGCCCAA 4534  
QY  
2550 CAA - - GGAATTCAAAGATGTTAAGGACTTAAAGCCCAAGGCTTAGTAAACCATGCA 2607  
Db  
4535 CAAAGGCTTCCAAAGATGTTAAGGACTTAAAGCCCAAGGCTTAGTAAACCATGCA 4594  
QY  
2608 TAACTCCCTGCAAGTAAATCCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGGACAGTGG 2667  
Db  
4595 TAACTCCCTGCAAGTAAATCCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGGACAGTGG 4654  
QY  
2668 AGGGTTAGTGAAGATCTCAGGATTAATCAATGGAGGCGTTCCTCTTTTATACCCAGCTG 2727  
Db  
4655 AGGGTTAGTGAAGATCTCAGGATTAATCAATGGAGGCGTTCCTCTTTTATACCCAGCTG 4714  
QY  
2728 TACCTAGCCCTTATCTGTMVTTCCCAATACAGAGGAACAGAGTGGTTTACASTCC 2787  
Db  
4715 TACCTAGCCCTTATCTGCTGCTTCCCAATACAGAGGAACAGAGTGGTTTACASTCC 4774  
QY  
2788 TGGACCTTMAAGATGCTTCTGCAATCCCTGTATCATCTGACCTCTCAATCTCTGTTG 2847  
Db  
4775 TGGACCTTMAAGATGCTTCTGCAATCCCTGTATCATCTGACCTCTCAATCTCTGTTG 4834  
QY  
2848 CTTTGTGAAGATCTTCAAAACCCACATCTCAACTCAGCTGGACTTTTACCCCAAGGCT 2907  
Db  
4835 CTTTGTGAAGATCTTCAAAACCCACATCTCAACTCAGCTGGACTTTTACCCCAAGGCT 4894  
QY  
2908 TCAGGGATAGYCCCATCTATTTGGCCAGGCAATAGCCCAAGACTTGGAGCTATTCAT 2967  
Db  
4895 TCAGGGATAGYCCCATCTATTTGGCCAGGCAATAGCCCAAGACTTGGAGCAATCTCAT 4954  
QY  
2968 ACTTGACACTCTGCTCTCTGCTACAGAGGTTAAATATCTTAGGCTAAATTA 3026  
Db  
4955 ACTTGACACTCTGCTCTCTGCTACAGAGGTTAAATATCTTAGGCTAAATTA 3086  
QY  
3027 ACTTGCTGCATCAAGCCACCAAGCCTCTTAAATTTCTCGYACCTGCTGCTACAG 3086  
Db  
5015 ACTTGCTGCATCAAGCCACCAAGCCTCTTAAATTTCTCGYACCTGCTGCTACAG 5074  
QY  
3087 GTTTTCAAAACSAARGCTCACTCTGCTACAGAGGTTAAATATCTTAGGCTAAATTA 3146  
Db  
5075 GTTTTCAAAACSAARGCTCACTCTGCTACAGAGGTTAAATATCTTAGGCTAAATTA 5134  
QY  
3147 TCCAAAGGCAACAGGCCCTCAGTGAAGAAATATCCAGCTTATCTGCTTATCTCAT 3206  
Db  
5135 TCCAAAGGTAACCAAGGCCCTGCTGAGGAAACGTATCCAGCTTATCTGCTTATCTCAT 5194  
QY  
3207 CYCAAAACCTTAAGCACTAAGRRTTCTTGGCTTAAGYAGGTTCTGCGCAATGAG 3266  
Db  
5195 CCAAAACCTTAAGCACTAAGRRTTCTTGGCTTAAGYAGGTTCTGCGCAATGAG 5254  
QY  
3267 ATTCCCAAGTGTGGCAATATAGCCAGGCTTAAATATCAATTAAGGAACTCAGAA 3326  
Db  
5255 ATTAA - CAGTATGGCAATATAGCCAGGCTTAAATATCAATTAAGGAACTCAGAA 5313  
QY  
3327 AGCCCAATACCCATTTATTAAGATGGAYAMCTGAAYMRAAGTGGCTTCCAGGCCCTTAA 3386  
Db  
5314 AGCCCAATACCCATTTATTAAGATGGAACTGAGCACTGAAGTAGAAGTGGCTTCCAGGCC - - - 5369  
QY  
3387 AGNAGGCTTAAACCAAGYCCAGTGTAAAGYTTGCGCAACGCGGCAAGACTTTTAA 3446  
Db  
5370 - - - - - TAAACCAAGCCCGTGTAAAGTGTGCGCAACGCGGCAAGACTTTTAA 5419  
QY  
3447 TAYRTACAGAAACCAAGAAATAGCTCTGAGGCTTACACAGRTCCAGGGAGAG 3506  
Db  
5420 TATGTCACAGAAA - AAACAGGAATAGCTCTAGGAGTCTTACACAGATCCGAGGATGAG 5478

QY 3507 CTTGCAACCVRTGGCRVACCTGASTAAGAAAYTGTATGTAGTGGCAAGGGTTGRCYTCA 3566  
Db 5479 CTTGCAACCTGTGGGTACCTGACTAAGAAATTTGTATGTAGTGGCAAGGGTTGRCYTCA 5538  
QY 3567 TTGTTTAYGGGTAGTGGTGGCAGTAGCAGTGYKTATCTGAAGCAGTAAATAATACA 3626  
Db 5539 TTGTTTAYGGGTAAATGGTGGCAGTAGCAATCTTAGTATCTGAAGCAGTAAATAATACA 5598  
QY 3627 GGGARAGATCTTACTGTGTGGAATCTCATGAKGTGAAYRGCACTACATCTGCTAAAGG 3686  
Db 5599 GGAAGAGATCTTACTGTGTGACATCTCATGATGTGAACAGCACTACTACTGCTAAAGG 5658  
QY 3687 AGACTTGTGGCTGTGACAGAAACGTTTACTTAAATRTCTAGGCTCTATTACTTTGAAGGCC 3746  
Db 5659 AGACTTGTGGCTGTGACAGCACTTTTACTTAAATACCAGGCTCTATTACTTTGAAGGCC 5718  
QY 3747 AGTGTGCRATGTGCACTTGTGCAACTTCTTAAACCCAGYCNCAATTTCTTCCAGACAATGA 3806  
Db 5719 AGTGTGCAACTGTGCACTTGGGCACTTCTTAAACCCAGCACTTTCTTCCAGACAATGA 5778  
QY 3807 AGAAAGATARAAYATACTGTCAACAARTAAATTTCTCAAACTATGCCACTCGAGGGA 3866  
Db 5779 AAAAAGATAGAAACGTAACCTGTCAACAGGTGATTGCTCAAACTATGCCCTCGAGGAA 5838  
QY 3867 CTTTGTAGAGCTTCCYTTGACTGATCCYGCCTTCAACTGTATATCTGATGGAAGTCTCT 3926  
Db 5839 CTTTGTAGAGCTTCCYTTGACTGATCCYGCCTTCAACTGTATATCTGATGGAAGTCTCT 5897  
QY 3927 TTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGCACTGCTGATGATGGAATATGTTG 3986  
Db 5898 TTGTAGAAAAAGGACTTCGAAAAAGCAGGATGCACTGCTGATGATGGAATATGTTG 5957  
QY 3987 AAAGTAATCCCTCACTCCAGGAACTAGTGTAGTGTGCAAGAACTAATAGCCYCTATK 4046  
Db 5958 AAATAATCCCTCACTCCAGGAACTAGTGTAGTGTGCAAGAACTAATAGCCYCTATK 6017  
QY 4047 GGGCACTAGAAATAGGAGAAAGGAAATATATATATATATATATATATATATATATAT 4106  
Db 6018 CGGCACTAAAAATAGAAAAAGGAAAGGTAACATATATATATATATATATATATATAT 6077  
QY 4107 ACCTAGTCTCCATCCCATGMRGCAATATATGSAAGAAAGGAAATTCCTAACTTCVAGR 4166  
Db 6078 ACCTAGTCTCCATCCCATGMRGCAATATATGSAAGAAAGGAAATTCCTAACTTCVAGR 6137  
QY 4167 GAACACCTTATCAMACATCAGGAAGCCATTAGGARATTTATTTATGCGWGTACAGAACTTA 4226  
Db 6138 GAACATCTTATCAAAACATCAGGAAGCCATCAGGAGATTTATTCGGCTGTACAGAACTTA 6197  
QY 4227 RAGAGTGTGMACTTTTACACTGCTGCTGCTCATCAAAAGGAAAGGAAAGGAAATASAG 4286  
Db 6198 RAGAGTGTGMACTTTTACACTGCTGCTGCTCATCAAAAGGAAAGGAAAGGAAATASAG 6257  
QY 4287 RGAAYTGCAAGCAKATATGAAGGMAAGAGCTGCAAGGCGAGGACCTCCATTAGAAA 4346  
Db 6258 GGAACCGCAAGCAAGATTTGAGCCAAAGAGCCACAGGCAAGGCAAGGCAAGGCAAGG 6317  
QY 4347 TGCCTTATTAAGCTTCCCTTAGTATAGGGAATTCCTT 4383  
Db 6318 TGCCTTATTAAGGACCCCTTAGTATAGGCAAGGCAAT 6354

## RESULT 9

US-10-450-763-12279  
; Sequence 12279, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631

;; PRIOR FILING DATE: 2001-03-30  
;; PRIOR APPLICATION NUMBER: 09/540,217  
;; PRIOR FILING DATE: 2000-03-31  
;; PRIOR APPLICATION NUMBER: 09/649,167  
;; PRIOR FILING DATE: 2000-08-23  
;; NUMBER OF SEQ ID NOS: 60736  
;; SOFTWARE: Custom  
;; SEQ ID NO 12279  
;; LENGTH: 4349  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: SIMILAR  
;; LOCATION: (466)..(906)  
;; OTHER INFORMATION: 100% homologous to Homo sapiens neurotensin  
;; OTHER INFORMATION: receptor, accession number X70070, Smith-Waterman Score=775.  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)..(4349)  
;; OTHER INFORMATION: n = a, t, c or g  
US-10-450-763-12279

Query Match 45.8%; Score 3431; DB 10; Length 4349;  
Best Local Similarity 87.3%; Pred. No. 0;  
Matches 3974; Conservative 148; Mismatches 195; Indels 237; Gaps 28;  
  
QY 115 ACTGCACCTCTTCGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCAC 174  
DB 1 ACTGCACCTCTTCGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCAC 60  
  
QY 175 CACTGCTGTTTGGCCACCACCGANACTCGCGCTGACTCCCATCCCTCTGGATCCCTGCAG 234  
DB 61 CACTGCTGTTTGGCCACCACCGAGACTGCGCTGACTCCCATCCCTCTGGATCCCTGCAG 120  
  
QY 235 GGTGTCGGCTGTCTCTGATCCAGCGA---RGCGCCATTTGCGCTCCCAATTTGGGCTA 291  
DB 121 GGTGTCGGCTGTCTCTGATCCAGCGAGCGCGGCCCATTTGCCATCTCCAGTCGGGCTA 180  
  
QY 292 AGGCTTGGCATTTGNCCTGAC- GGTAAAGTGCCT- GGGTTTGTCTTAATTGAGCTGAA 349  
DB 181 GAGGCTCACCATTTGTTCTGACAGCGCTAAGTGCCTGGGGTTTCATCTAATCGAGCTGAA 240  
  
QY 350 CACTANTCACTGGTTCATGTTCTCTCTGACCGCCAGCGCTTCTTAATAAATTAATTA 409  
DB 241 CACTAATCACTGGTTCATGTTCTCTCTGACCGCCAGCGCTTCTTAATAAATTAATTA 300  
  
QY 410 CACTTACCACATGGCCCAAGATTCCATTTCTTGGAAATCCGTGAGGSCAACGAACTCCAG 469  
DB 301 CACTTACCACATGGCCCAAGATTCCATTTCTTGGAAATCCGTGAGGSCAAC- GAATCTCAGG 359  
  
QY 470 TCAGAGAAATCGAGCTTGCCACCATCTTTGGAAGCGGCTGTACCTCTTTGGAAGTGGT 529  
DB 360 TCAGAGAAATCGAGCTTGCCACCATCTTTGGAAGCGGCTGTACCTCTTTGGAAGTGGT 419  
  
QY 530 TCACCAACCATCTTGGAGCTCTGTGACGAGACGACCCCGGTGACATTTTGGCRACGMS 589  
DB 420 TCACCAACCATCTTGGAGCTCTGTGACGAGGACCCCGGTGACATTTTGGCAACCAAC 479  
  
QY 590 RACGACATCTCMAAGTGTGGGAAACGTTTCCCGCAAGACAAACCGCCCTTAAGACGTA 649  
DB 480 GACGGACATCCAAAGTGTGGGAAACGTTTCCCGCAAGACAAACCGCCCTTAAGACGTA 539  
  
QY 650 TTCTGGAAATTTGGGAAATTTGACCTCTCAGACACTAAGAAAGAAACGACTTATATCT 709  
DB 540 TTCTGGAAATTTGGGAAATTTGACCTCTCAGACACTAAGAAAGAAACGACTTATATCT 599  
  
QY 710 TCTGAGTGGCGCTGGGACTCTCTGAGGGAATTAATAATTAACACCATCTTTACAGCTA 769  
DB 600 TCTGAGTGGCGCTGGGACTCTCTGAGGGAATTAATAATTAACACCATCTTTACAGCTA 659  
  
QY 770 GACTCTTTTGTAGAAAGGCAATGGAGTGAAGTGCCTAAGTACAACTTTCTTTTCA 829  
DB 660 GACCTCTTTTGTAGAAAGGCAATGGAGTGAAGTGCCTAAGTACAACTTTCTTTTCA 719

QY 830 TTAAGAGACAACTCAAAATTTATCTAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAG 889  
DB 720 TTAAGAGACAACTCAAAATTTATCTAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAG 779  
  
QY 890 AGTCTACCTCCCTATCCAGCATCCCGACTCTTCCCAATTAAGGACCCCTTC 949  
DB 780 AGTCTACCTCCCTATCCAGCATCCCGACTCTTCCCAATTAAGGACCCCTTC 839  
  
QY 950 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACAAAGAGTGCAC 1009  
DB 840 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACAAAGAGTGCAC 899  
  
QY 1010 TATTCGCCAATTAAGACCTCCCAAGCAGTGGAGGAGAGAAATTCGGCCACGACAG 1069  
DB 900 TATTCGCCAATTAAGACCTCCCAAGCAGTGGAGGAGAGAAATTCGGCCACGACAG 958  
  
QY 1070 TGCATGTGCTTTTCTCCAGCATTTAAAGCAAAATTAAGACAGACTTAGGTAAATCT 1129  
DB 959 TGCATGTGCTTTTCTCCAGCATTTAAAGCAAAATTAAGACAGACTTAGGTAAATCT 1018  
  
QY 1130 CAGATAAATCTGATGCTATATTTGTTTCAAGGTTTAGGACAAATCTTTGATCTGA 1189  
DB 1019 CAGATAAATCTGATGCTATATTTGTTTCAAGGTTTAGGACAAATCTTTGATCTGA 1078  
  
QY 1190 CATGGAGAGATATATGTCATGCTTAAATCAGACACTAACCCCAATGAGAGAGTGCC 1249  
DB 1079 CATGGAGAGATATA-ATGTCATGCTTAAATCAGACACTAACCCCAATGAGAGAGTGCC 1137  
  
QY 1250 ACCATAAATGTCAGCTGAGGTTTGGGATCTCTGATCTCACTGAGTCAATGGATAN 1309  
DB 1138 ACCATAAATGTCAGCTGAGGTTTGGG- TCTCTGATCTCTCACTGAGTCAATGGATAN 1196  
  
QY 1310 GGATGACAAAGAGAAAGAAATGATTTCCACAGGCGCCAGCAGCTTCCAGCTA 1369  
DB 1197 GGATGACAAAGAGAAAGAAATGATTTCCACAGGCGCCAGCAGCTTCCAGCTA 1256  
  
QY 1370 SACCCTCATTTGGGACACAGAAATCAGTAACATGGGAGATTTGCTGTCAGACATTTGCT 1429  
DB 1257 GACCCTCATTTGGG--ACACAGAAATCAGAACAT- GGAGATTTGCTGTCAGACATTTGCT 1312  
  
QY 1430 AACTTGTGCTGCTASAGGACTAAGGAAACTAGAGAAATCTAGAGAAATTAATTAATGA 1489  
DB 1313 AACTTGTGCTGCTAGAGGACTAAGGAAACTAGGAAAG- AAGTCTATGAATTAATTAATGA 1371  
  
QY 1490 TGTCCACCATTAACAGAGGAGGAGAAATCTCTACTGCTTCTGAGGAGACTAAGG 1549  
DB 1372 TGTCCACCATTAACACA- GGGAGAGGAGAAATCTCTACTGCTTCTGAGGAGACTAAGG 1430  
  
QY 1550 GAGGCATTTGAGGAAGCTGCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTTAATCTTA 1609  
DB 1431 GAGGCATTTGAGGAAGCTGCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTTAATCTTA 1490  
  
QY 1610 AAGCGTAAGTTTATCACTCAGTGCAGTGCAGATTTAGAAAAAACTTCAAAAGTCTGCCG 1669  
DB 1491 AAGCGTAAGTTTATCACTCAGTGCAGTGCAGATTTAGAAAAAACTTCAAAAGTCTGCCG 1550  
  
QY 1670 TAGGCGCGGAGCAAACTTAGAAACCTTATGAACTTGGCAACTCGGTTTTTATTAATA 1729  
DB 1551 TAGGCGCGGAGCAAACTTAGAAACCTTATGAACTTGGCAACTCGGTTTTTATTAATA 1610  
  
QY 1730 GAGATCAGGAGGAGCAGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTA 1789  
DB 1611 GAGATCAGGAGGAGCAGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTA 1670  
  
QY 1790 GTCTAGACCTCAGCAAGTGGACTTTGGAGCTCTGGAAGGAGGAAAAAGCTGGGCAAT 1849  
DB 1671 GTCTAGACCTCAGCAAGTGGACTTTGGAGCTCTGGAAGGAGGAAAAAGCTGGGCAAT 1730  
  
QY 1850 TGAATGCCCTAATAGGCTTGTTCAGTGGCGTCTCAAGGACACTTTAAAAAAGATTTGT 1909  
DB 1731 TGAATGCCCTAATAGGCTTGTTCAGTGGCGTCTCAAGGACACTTTAAAAAAGATTTGT 1790



1910 CCAAGTAGAAGTAAGCGGCCCTTCGTCCATGCCCTTATTTC AAGGAATCACTGGAAG 1969  
1791 CCAAGTAGAA----- 1800  
1970 GCCCACTGCCCCAGGGGACAAAGGTCTTTTGTAGTCAGAACCACTAACACAGATGATCCAG 2029  
1801----- 1800  
2030 CAGCAGGACTGAGGGTGCCCTGGGGCAGCGCCATCCCATGCCATCACCCCTACACAGGCC 2089  
1801-----AGCC 1805  
2090 TGGGTATGCTTACACATTCAGAGCGCCAGGAAGT---TGTCCTCTGGACACTGTCGGTTC 2146  
1806 CAGGTATGCTTGACCATTTAGAGCCCAAGAGGTTAAGTGTCTCTGGACCGCGGTGCGGCC 1865  
2147 TTCTTAGTCTTACTCTCTCTGTCGCCGACAACTGTCTCCAGATCTGTCACTATCC--GA 1922  
1866 TTCTCAGTCTTACTCTCTGTCGCCGACAACTGTCTCCAGATCTGTCACTATCC--GA 1922  
2207 GGGTCNTAAGACGGGCACTCAGATACATCTTTTTCCTGAGCCCACTAAGTTATGAACCTGGG 2266  
1923 GGGGCTCTAGGACAGCCACTCAGATAC--TTCTTCCAGCCACTAAGTTGTG-ACTGGG 1980  
2267 GAGCTTATTTTTCACATGCTTTTCTAATTAATCTTGAAAGCCCACTACTCTGTTAG 2326  
1981 GAGCTTTTACCTTTTTCACATGCTTTCTTAATTAATGCTTGAAAGCCCTCACTCCCTTGTAG 2040  
2327 GGAGAGACATTTAGCAAAAGCAGGGGCCATTTATACCTGAACATAGAGAGGAACAC 2386  
2041 AGAGAGAGATTTAGTAAAGCAGGGCCATTTATACCTGAACATAGAGAGGAACAC 2100  
2387 CGGTTTGTGTCCTCTGTCAGGAGGAATTAATTCCTGAAAGTCTGGGCAACAGAGGA 2446  
2101 CCAATTTGTCGT-CTCCTGTTTGGAGGAAGGAATTAGTCTGGAAGTCTGGGCAACAGAGGA 2159  
2447 CAATATGGAGGACCAAGAAATGCCCTCTGTTCAAGTTAACTAAAGGATTCACATTC 2506  
2160 CAATATGGAGCAAG-CAAGAAATGCCCATTTTGTCCAAAGTTAACTAAAGGATTCACATTC 2218  
2507 CTTTCCCTTACCAAGGCACTACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAGAT 2566  
2219 CTTTCCCTTACCAAGGCACTACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAGAT 2278  
2567 TGTTAAGGACTTAAAGCCAGGCTTAGTAAACCAATGCAATACCTCCCTGCAAGTAATTC 2626  
2279 TGTTAAGGACTTAAAGCCAGGCTTAGTAAACCAATGCAATACCTCCCTGCAATACCTCC 2338  
2627 CGTAGTGGATTGAGGAGGACAGAACCCAGTGGAGGAGGTTAGTGCAGATCTC 2686  
2339-----AATTTTAGGAGTACAGAAACCCAAAGACAGTGGGA-GGTTAGTGCAGATCTC 2390  
2687 AGGATTAATCAATGGAGCGGTGTCTCTTTTATACCCAGCTGTACCTAGGCCCTTATACCTGT 2746  
2391 AGGATTAATCAAT-GAGGCTGTTTTCTCTATACCCAGCTATACCTAGGCCCTTATACCT 2449  
2747 GMYTTCCCAAAACACAGAGGAGCAGAGTGGTTTACATCTCTGGACCTTMAAGATGCCCTT 2806  
2450 GCTTTCCCAAAATACACAGAGGAGCAGAGTGGTTTACAGTCTCTGGATCTTAAAGGATGCCCTT 2509  
2807 CTTCTGCATCCCTGTACATCTGACTCTCAATTTCTGTTGCTTTGAGATACCTTCAA 2866  
2510 TTTCTGCATCCCTGTACATCTGACTCTCAATTTCTGTTGCTTTGAGATACCTTCAA 2569  
2867 CCARCATCTCAACTCACTGACTTTTATCCCAAGGGTTCCAGGATAGYCCCCTCT 2926  
2570 CCCAATCTCAACTCACTGACTTTTATCCCAAGGGTTCCAGGATAGTCCCATCT 2629  
2927 ATTTGGCCAGGATTAAGCCAGACTTGAGYCAATYMTATACCTGGAACACTTGTCT 2986  
2630 ATTTGGCCAGGATTAAGCCAGACTTGAGCCAGTCTCATACCTGGAACACTCTGTCTCT 2689  
2987 TCRGTAKGTGGATTTACTTTTRCGYCCYRTTTCAGAAACCTTGTGCGCCATCAGCCAC 3046

2690 TCAGTGCATGATGATTTACTTTTAGCTGCCGTTTCAGAAACCTTGTGCGCATCAAGCCAC 2749  
3047 CCAAGCRCTCTTMAATTTCTCGCYACCTGTGGCTACAGWGTTCCTCAAAACRSARAGCTCA 3106  
2750 CCAAGCGCTCTTAAATTTCTCTACCTACCTGTGGCTTACAAGGTTTCCAAACCAAGGCTCA 2809  
3107 RCTGTCTCACAGCAGGTTAAATATTAGGCTTAAATTTATCCAAAGCACCARGGCCCT 3166  
2810 GCTCTGCTCACAGCAGGTTAAATGCTTAGGGCTTAAATTTATCCAAAGTACCARGGCCCT 2869  
3167 CAGTGAGGAAYRYATCCAGCCTTATCTGCTTATCTCATCYCAAAACCTTAAAGCAACT 3226  
2870 CAGTGAGGAACGTATCCAGCCTTATCTGCTTATCTCTTATCCAAACCTTAAAGCAACT 2929  
3227 AAGRRRTTCTTGGCCTTAAAYAGGTTCTGCGAATATGGAATCCCAAGGTTGCGRAAA 3286  
2930 AAGGGTTCCTTGGCATAACAGGTTTCTGCGAATATGGAAT-CCCAGGTACGCGCAAA 2988  
3287 TAGCCAGGYCATTAATATACASTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTARTAA 3346  
2989 TAGCCAGACCGTATATACGCTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTAGTAA 3048  
3347 GATGAYAMCTGAAGYMAAGTGGCTTTCAGGCCCTTAAAGAGGCTTAAACCCCAAGY 3406  
3049 GATGACACCTGAAGCAGACGCTTTCATGCCCTTAAAGAGGCCCT--NATGCAAGC 3106  
3407 CCCAGTGTAAAGYTTGCAACRGGCCAAGACTTTTSTYATATYRTCA- 3454  
3107 CCCAGTGTAAAGTGTCTAAGGGGCAAGACTTTTCTCATATGTCACAAAAAACAAACA 3166  
3455-----AGAAAAACAGAAACATGAATAGCTCTAAGAGTCTTGCGTGGTCTGAGGACGA 3505  
3167 AACAAAGAAAAAATGAATAGCTCTGAGTCTTACACAGRTCCBAGGAGCA 3505  
3506 GCTTCCAAACCYRTGGCRVACCTGAAGGAAAYTGATGTAGTGCAAAAGGGTTGRCYTC 3565  
3227 GTTTCGAACCGTGCATACCTGA--TAAGGAATTTGATGTAGTGCAAAAGGGTCAGCCTC 3285  
3566 ATTTGTTTAYGGTGTGTGGCAGTACGAGTGTAGTATCTGAAAGAGTTAAATAATAC 3625  
3286 ATTTGTTTACGGGTAGTGTGGCAGTACGAGTCTAGTATCTGAAGCAGTTAAATAATAC 3345  
3626 AGGPRAGAGATCTTACTGTGGACATCTCATGAGTGAAYRGCTACTCCTGCTTAAAG 3685  
3346 AGGGAAGAAATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCCTGCTTAAAG 3405  
3686 GAGACTTGTGGTGTGCAGACAACTGTTTACTTAAATRTCAGGCTCTATTACTTGAARGC 3745  
3406 GAGACTTGTGGTGTGCAGACAACTTACTTAAATRTCAGGCTCTATTACTTGAARGC 3465  
3746 CAGTGTCTGACTGTGCACCTGTGCAACTTTAACCCAGYCNACTTTCTTCAGACAAATG 3805  
3466 CAATCTCGCACTGTGCACCTGTGCGAGCTTTAACTGACACATTTCTTCCACAAATG 3525  
3806 AAGAAAGATBAAAYATTAACCTGCAACAAATTAATTTCTCAAACTTATGCACTCGAGGG 3865  
3526 AAGAAAGATBAAACATACTGTCAACAAATTAATTTGCTCAAACTTATGCACTCGAGGG 3585  
3866 ACCTTGTAGAGTTCCTTGTACTGTATCCVAGCTTCAACTTGTATATCTGATGGAAGTCC 3925  
3586 ACCATTTAGAGTTCCTTGTACTGTATCCCACTTACTGATGATGGAAGTCC 3644  
3926 TTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGAGTGGTTCAGTGAATGAATATYTT 3985  
3645 TTTGTAGAAAAAGGACTTCGAAAAAGTGGGTATGAGTGGTTCAGTGAATGAATATYTT 3704  
3986 GAAAGTAAATCCCTCACTCCAGGAATAGTGTCTGAGTGGCAGAACTAATAGCCYCAVT 4045  
3705 GAAAGTAAATCCCTCACTCCAGGAATAGTGTCTGAGTGGCAGAACTAATAGCCYCAVT 3755  
4046 KGGGCACTAGAAATTAAGGAGAAAGGGAATATATATACAGACTCTTARTATGCT 4105



```
Db 3756 CGGCACTAGATTAGGAGGAGAAAGGGTAATATATATACAGACTCTAAGTATGCT 3815
Qy 4106 YACCTAGTCNTCCATGCGCCATGMRGCAATATGTSARAGAAAGGGAATTCCTAACTTCYAG 4165
Db 3816 TACCTAGTCCTCCATGCGCCATGCAAGTAAATAGGAGAGAAAGGGAATTCCTAACTTCYAG 3875
Qy 4166 RGAACACTATCAMACACTCAGGAGCCATTAGGARATTTATVATGCGWGTACAGAAACCT 4225
Db 3876 GGAACACTATCAACACTCAGGAAGCCATTAGGGAATTTATVATGCGCGTACAGAAACCT 3935
Qy 4226 ARAGAGTGMAGTCTTACACTGCGGGGTCAATCANAAAGGAAAGRAAGAAAGGGAATASAA 4285
Db 3936 AAAGAGTGGCAGTCTTACACTGCGGGGTCAATCAGAAAGGAAAGGGAATAGAA 3995
Qy 4286 GRGAAYGCCAAGCAKATATTGAAGMAAAGAGCTGCAAGCGCAGGACCTCCATTTAGAA 4345
Db 3996 AGGAACCGCCAGCAGATATTGAAGCCAAAGAGCTGCAAGCGCAGGACCTCCATTTAGAA 4055
Qy 4346 ATGCTTATTAACCTTCCCTTAGTATAGGGTAAATCCCTTCCGGGAAACCAAGCCCCAGTAC 4405
Db 4056 ATGCTTATAGAGGACCCCTTAGTATAGGGTAAATCCCTTCCGGGAAACCAAGCCCCAGTAT 4115
Qy 4406 TCAGCAGGAGAAACAGAAATGGGAACCTCACGAGG---CAGTTTCTCCCTCGGAGCGG 4462
Db 4116 TCAGNAGAGAAATAGAAATGGGAATCTCAGGAGCATAGTTTCTCCCTCAGGATGG 4175
Qy 4463 TTAGCCACTGAAGAGGAAATATCTTTGCGCTGCAACTATCCAATGGAAATTAATTTAAA 4522
Db 4176 CAAGCACCAGAAAGGAAATACCTTTGCTGCACTAACCAATGGAATTAATTTAAA 4235
Qy 4523 ACCCTTCATCAACCTTTCACTTAGGCATCGATAGCAGCCCATCARATGGGCAATTTA 4582
Db 4236 ACCACTCACCAGAACTTTCACTTATGATGATAGCAGCCCATCAGATGGGCAATTTGTTA 4295
Qy 4583 TTTACTGGACAGGCTTTTCAAACTATCAAGCARATATKTCAGGGCTCTGTAA 4636
Db 4296 TTTACTGGACAGGCTTTTCAAACTCTCAAGCAGATAGTTGGGCTCTGTAA 4349

RESULT 10
US-10-450-763-4430
; Sequence 4430, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 4430
; LENGTH: 7466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1378)..(2271)
; OTHER INFORMATION: 84% homologous to multiple sclerosis associated retrovirus
; OTHER INFORMATION: polyprotein, accession number AF009668, Smith-Waterman Score=1279.
US-10-450-763-4430

Query Match 44.2%; Score 3309.6; DB 10; Length 7466;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 3417; Conservative 98; Mismatches 17; Indels 38; Gaps 12;
Qy 1 CAACAATCGGGATATAAACCAGGCATTCGAGCTGGCAACAGCAGCGCCCTTTGGGTGC 60
```

```
Db 3924 CAATGATCGGGATATAAACCAGGCATTCGAGCTGGCAACAGCAGCGCCCTTTGGGTCC 3983
Qy 61 CTTCCCTTTCTATGGGAG--CTGTTTTCTATGCTATTTCTCTCTATTAATCTTGCACACTG 118
Db 3984 CTTCCCTTTCTATGGGAGCTCTGTTTTCTATGCTATTTCTCTCTATTAATCTTGCACACTG 4043
Qy 119 CACTCTTCGGTCCATGTTTCTTACCGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACT 178
Db 4044 CACTCTTCGGTCCATGTTTCTTACCGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACT 4103
Qy 179 GCTGTTTGGCCACACCGCANACCTGCGGCTGACTCCCATCCCTCTGATCTGATCTGAGGGTG 238
Db 4104 GCTGTTTGGCCACACCGCAGACCTGCGGCTGACTCCCATCCCTCTGATCTGATCTGAGGGTG 4163
Qy 239 TCCGCTGTGCTCCTGATCCAGCGGARGCGCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTT 298
Db 4164 TCCGCTGTGCTCCTGATCCAGCGGAGCGGCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTT 4223
Qy 299 GCCATTGTGCTGCAAGGCTAAGTGTGCTGGGTTGTTCTAATTGAGCTGAACACTANTCA 358
Db 4224 GCCATTGTGCTGCAAGGCTAAGTGTGCTGGGTTGTTCTAATTGAGCTGAACACTANTCA 4283
Qy 359 CTGGGTTCCATGTTTCTCTCTGCTGACCCACGCGTCTTAAT-----AKAACTATAA 409
Db 4284 CTGGGTTCCATGTTTCTCTCTGCTGACCCACGCGTCTTAATAGAGCTTAAACAGAGCTATAA 4343
Qy 410 CACTTACCATCATCGGCCCAAGATTCCATTCCTCTGTCGACCCACGCGTCTTAATAGAGCTATAA 4343
Db 4344 CACTTACCGCATGCGGCCAAGGTTCCGTTCTTGTAAATCCGTCGAGGCCAA--GAACCCCGAG 4402
Qy 470 TCAGAGAATACGARGCTTGCACCATCTTTGGAAAGCGGCTGTACCTCTTTGGAAGTGGT 529
Db 4403 TCAGAGAACAAAAGGCTTGCACCATCTTTGGAAGCGGCTGTACCTCTTTGGAAGTGGT 4462
Qy 530 TCACCAACCATCTTTGGGAGCTCTGTGAGCAAGGACCCCGCGGTACATTTTGGGACCAAMS 589
Db 4463 TCACCAACCATCTTTGGGAGCTCTGTGAGCAAGGACCCCGCGGTAACTTTTGGCAACCAAC 4522
Qy 590 RACGGACATCCMAAGTATGGGAAACGTTCCCGCAAGACAAAAGCCGCTTAAGAGCTA 649
Db 4523 GACGGACATCCAAAGTATGGGAAACGTTCCCGCAAGACAAAAGCCGCTTAAGAGCTA 4582
Qy 650 TTCTGGARAATTTGGGAMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTTATATCT 709
Db 4583 TTCTGGARAATTTGGGACCAATTTGACCTCAGACACTAAGAAAGAAACGACTTTATATCT 4642
Qy 710 TCTGAGTCCGCTCTGGCACTCTCTGAGGAAAGTATAAATTAACACCATCTTTACAGCTA 769
Db 4643 TCTGAGTCCGCTCTGGCACTCTCTGAGGAAAGTATAAATTAACACCATCTTTACAGCTA 4702
Qy 770 GACTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCA 829
Db 4703 GACTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCA 4762
Qy 830 TTAAGAGACAACTCACAAATTAATAAAGTGTGATTTATGCCCTACAGAGACCTTCAG 889
Db 4763 TTAAGAGACAACTCACAAATTAATAAAGTGTGATTTATGCCCTACAGAGACCTTCAG 4822
Qy 890 AGTCTACCTCCCTATCCAGCATCCCGACTCCTTCCCAATTAATAAGGACCCCTTTC 949
Db 4823 AGTCTACCTCCCTATCCAGCATCCCGACTCCTTCCCAATTAATAAGGACCCCTTTC 4882
Qy 950 RACCCAAATGTCCTAAAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCAA 1009
Db 4883 AACCCAAATGTCCTAAAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCAA 4942
Qy 1010 TATTCCTCCCAATTAATGACCTCCAGCAGTGGGAGAGAGAAATTCGCGCCAGCCAGAG 1069
Db 4943 TATTCCTCCCAATTAATGACCTCCAGCAGTGGGAGAGAGAAATTCGCGCCAGCCAGAG 5002
Qy 1070 TGCATGTGCTTTTCTTCCAGACTTAAAGCAATAAATAAAGAGACTTAAAGTAAATCTT 1129
```



```

3290 CCAGCYCATTATWATACASTAAITTAAGGAAACTCAGAAAGCCAAATACCCATTTTARTTAAGAT 3349
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7212 CCAGTGCATTAATATACACTAATTAAGGAACTCAGAAAGCCAAATACCCATTTAGTAAGAT 7271
Qy      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
3350 GGAYAMCTGAAGYMAAGTGGCTTTCCAGGCCCTTAAAGAAGGCCTTAAACCCCAAGYCCC 3409
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7272 GGACAACCTGAAGTAGAGTGGCTTTCCAGGCC-----TAAACCAAGCCCC 7317
Qy      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
3410 AGTGTTAAGYTTGCCAACRGGGCAAGACATTTTSTTATATYRTCACAGAAAAAAACAGRAA 3469
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7318 AGTGTTAAGTGTGCCAACAGGGCAAGACTTTCTTCATATGTCACAG-AAAAAACAGGAA 7376
Qy      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
3470 YAGCTCTGGAGTCTCTACACAGRTCCRAGGAGYAGCTTGCAACCYRTGGCRYACCTGA 3529
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7377 TAGCTCTAGGAGTCTCTACACAGATCCGAGGATGAGCTTGCAACCTGTGGGTACCTGA 7436
Qy      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
3530 STAAGGAAAYTGATGTAGTGCCAAAGGTT 3559
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7437 CTAGGAAATYGATGTAGTGCCAAAGGTT 7466

RESULT 11
US-11-121-086-79
; Sequence 79, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIORITY FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79
; .LENGTH: 177175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-79

```

Query Match	41.5%	Score 3105.4	DB 15	Length 1717175
Best Local Similarity	86.7%	Pred. No. 0		
Matches 3580	Conservative 146	Mismatches 313	Indels 88	Gaps 29
QY	606	GATGGGAAACGTTCCCGCAAGACAAAAACCCCTAGACGTATTTCTGGARAATTGGGA	665	
Db	140106	GATGGGAATGTTTCCCCCAAGGCGAANAATGCCCTAAGATGTATTTCTGGAGAAATGGGA	140165	
QY	666	MCAATTTGACCTCAGACACCTAAGAAAGAAACGACTTATATTTCTTCGAGTGCCGCGCTG	725	
Db	140166	CCAATCTGACCATCAGACACCTAAGAAAGAAATGACTTATATTTCTTCGAGTACCACCTG	140225	
QY	726	GC-----ACTCTGAGGGAAGTATAAATTATAA	753	
Db	140226	GCCAATAATCTTCTCAAGGGGCGAGAACTCGGCTCTCGAGGGAAGTATAAATTATAA	140285	
QY	754	CACCATCTTAGAGCTAGACACTTTTGTAGAA---AAGGCAAAATGCGAGTGAAGTGCCATA	810	
Db	140286	CACCATCTTAGAGCTAGACCTCTTTTGTAGAAAAGGCAATGGAGTGAAGTGCCATA	140345	
QY	811	AGTCAAACTTTTTCATTTAAGAGACAACTCAAAATTATGTAAAAAGTGATTTATG	870	
Db	140346	TGTAACAACTTTTTCATTTAAGAGATAAATCCCAATTATGTAAAAAGTGATTTATG	140405	
QY	871	CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCAGCA--TCGCCGACTCCTTCCCC	928	
Db	140406	CCCTACAGGAAGCCTTCAGAGTCTACCTCCCGACCCCAAGACACCCCACTCTCTTCTCC	140465	
QY	929	AMYTAAATAGGACCCCGCTTCAACCCCAAAATGGTCCAAAGGAGATAGACAAAGGGTAAA	988	
Db	140466	AACTAATTAAGGACCCCGCTTCAACCCCAAAATGGTCCAAAGGAGATAGACAAAGGGGTAAA	140525	

QY	989	CAGTGAACAAAGAGTGCCAATATTTCCCAATATTTGACCCCTCCCAAGCAGTGGGAGGAA	1048
DB	140526	CAATGAACAAAGAGTGCCAATATTTACAGATATTTACTCGCTCC-----AAGCAGTGGGAG	140581
QY	1049	GAGAAATTCGGCCACGCCAGAGTGCATGTCTTTTYYTCTCCAGACTTTAAAGCAAAATAA	1108
DB	140582	GNGAATTTGGCCACGCCAGGTGCATGTACCTTTTCTCTCAGATTTAAAGCAAAATTA	140641
QY	1109	AAACAGACTTAGTAAATTTCTCAGATAAYCCTGATGGCTATATTGRTGTTTTTACAAGGGT	1168
DB	140642	AAATAGACTTAGTAAATTTCTCAGATAAACCTGATGGCTATATTGATGTTTTTACAAGGGT	140701
QY	1169	TAGGCAATTTCTTGATCTGACATGGAGAGATATATATGTCTACTGCTGTAATCAGACACTA	1228
DB	140702	TAGGCAATTCCTTTGATCTGACATGGAGAGATATA-ATGTACTCTGCTAAATCAGACACTA	140760
QY	1229	ACCCCAATGAGAGAAAGTGCACCACTAATCTCAGCTCAGAGTTTGGCGACTCTCTGGTAT	1288
DB	140761	ACCCCAATGAAAAAGTGTGCCATAACAGCAGCTCAGAGTTTGGCGAACTCTGGTAT	140820
QY	1289	CTCAGTCAAGTCAATTGGATANGGATGACACAGAAAGGAAAGANAATGATTCGCCACAGGC	1348
DB	140821	CTCAGTCAAGTCAATG--ATAGGATGACAAAG-ATGAAGAGATGATTTCCCAACAGGC	140877
QY	1349	CAGCARGCAGTTCCAGTCTTASACCTCATTTGGGACACAGAAATCAGTAACATGGGAGA	1408
DB	140878	CAGCAGGCAGTTCCAGTGTAGACCTCATTT-AGGACACAGAA--TCAGAACTTTGGAGA	140933
QY	1409	TTGGTCTGACAGACATTTGCTTAACTTTGTGCTASAAAGACTTAAGGAAAACTASGAAGNA	1468
DB	140934	TTGGTGCCACAGACATTTGCTTAACTTTGCTGCTAGAAGGACTTAAGGAAAACTTAGGAAG-A	140992
QY	1469	ARTCTAYGAAATTACTCAATGATGTCCACATTAACACAGGGGAAGGGAAGAAATCCTACT	1528
DB	140993	AGCCATGAATATTCAATGATGTCCCCTATPAACACA--GGGAAGGAAGAAATCCTACT	141051
QY	1529	GCCTTTCTGGAGAGACTAAGGGAGGCAATTGAGGAAGCGTGCTCTCTCTCACTGACTCT	1588
DB	141052	GCCTTTCTGGAGAGACTAAGGGGAAGGATTGAGGAAGCATACCTCCCTGTCACTGACTCT	141111
QY	1589	TCTGAAGGCCAACTTAATCTTAAAGGTTAAGTTTATCACTCAGTCAGCTGCAGACATT-AG	1647
DB	141112	ATTAAAGGCCAACTTAATCTTAAAGGATAAGTTTATCACTCAGTCAGCTGCAGAGATTAA	141171
QY	1648	AAAAAACTTTCAAAAGTCTGCGTAGGCCCGGAGCAAACTTTAGAAAAACCTATTGAACATTG	1707
DB	141172	AAAAAACTTTCAAAAGTAGTGCCTTAGGCCAGACAAAACTTTAGAAAAACCTACTGAACATTG	141231
QY	1708	GCAACYTGGTTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGAT	1767
DB	141232	GCAACCTCAGTTTTTTTTTATATAGAGATCAGGAAGAGCAGG--GGAATGGGACAAATGGAT	141290
QY	1768	-----TAAAAAAAGCCACCGCTTTAGTCATGACCTCAGGCAAGTGGACTTTGGAG	1820
DB	141291	AAAAAAAGGAAAGAGCTGGGCAAAATGAAATGCCTTAATAGGCTTGCTTCCAGTCGG	1880
QY	1821	GCTCTGAAAGGAAAGAGCTGGGCAAAATGAAATGCCTTAATAGGCTTGCTTCCAGTCGG	1940
DB	141351	GCTCCAGAAAGGAAAGAGCTGAGCAAAATGAAATGCCTTAACAGGCTTGCTTCTAGTGTG	141410
QY	1881	GTCTACAAGGACACTTTTAAAAAGATTGTCCAGTAGAAGTAAAGCGCGCCCTTCGTCCAT	1940
DB	141411	GTCTACAAGGACACTTTTAAAAAGATTGTCCAGTAGAAGTAAAGCGCTTGCTTCTAGTGTG	141470
QY	1941	GCCCTTTATTTCAAGGAAATCACTTGAAGGCCCACTGCCCCAGGGGACAAAGGCTTTTGTG	2000
DB	141471	GCCCTTTATGTCAAGGAAATCACTTGAAGGCCCACTGCCCCAGGAGATGAAGGCTCTCTG	141530
QY	2001	AGTCAGAAGCCACTTAAACAGATGATCAGCAGCAGGACTGAGGCTGCTTGGGGCAAGGC	2060
DB	141531	AGTCAGAAGCCACTTAAACAGATATCAGCAGCAGGACTGAGGATGCCAGGGCAAGGC	141590

QY	2061	CATCCATGCCATCAACCCCTCAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAG	2120	Db	142655	CTAAAAATTATCCAAAGGCTTGAGGGCCCTCAGTGAGGAACACATCCAGCCTATATCTGGCT	142714
Db	141591	CAGCCCATGCCATCAACCCCTCAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAG	141650				
QY	2121	GT--TGTCTCTCTGGACACTGGTGGGTCTTCTTATAGTCTTACTCTTCTGTCCTGGACAAAC	2177	QY	3198	TATCTCTCATCYCAAAAACCTTAAGCAACTAGRGRPTTCTTGGCTTAAYAGGYTTCTGC	3257
Db	141651	TTCAGTGTCTCTGGACACTGGTATGGCTTCTCAGTCTTACTCTCTCTGTCCTGGACAAAC	141710	Db	142715	TATCTCTCTCCAAAACCTTAAGCAACTAGAGGGTTCCTTGGCATTAACAGGTTCTTAC	142774
QY	2178	TGTCTCTCAGATCTGCTCACTATTCTGAGGGGTCTNTAAGCGGCGAGCTCACTAGATACT	2237	QY	3258	CGAAWATGGATTCCCCAGGTGTGCRAAATAGCCAGGYCATTTAWATACASTAAATTAAGA	3317
Db	141711	TGTCTCTCAGATCTGCTCACTATCC--GAGGGGTCTTAGGACAGCAGCTCACTAGATAC-	141766	Db	142775	CGAATGTGGATT-CCAGGTACAGCAAAATAGGCAGATCATTTATATACACTAAATTAAGA	142833
QY	2238	TTTCTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTTTACATGCTTTTCTAAT	2297	QY	3318	AACTCAGAAAACCCCAATACCCATTTTARTAAAGATGGAYMCTGAAGYMRAGTGGCTTTCCA	3377
Db	141767	TTCTCCAGCCACTAGTTGTG-ACTGGGAACCTTTACTCTTTTTCACATGCTTTTCTAAT	141825	Db	142834	AACTCAGAAAACCCCAATACCCATTTTARTAAAGATGGACACCTGAAGCAGAGCAGCTTTCCA	142893
QY	2298	TATGTTTGAAGCCCCCACTACCTTTTATAGGGAGAGACATTTCTAGCAAAAGCAGGGGCCAT	2357	QY	3378	GGCCCTTAAGAAGGCCCTTAACCCAGYCCAGTGTGTTAAGYTTGGCCAAACRGGGCAAGAC	3437
Db	141826	TATGGTGAAGGCCCACTCCCTTGTAGGGAGAGACATTTCTAGCAAAAGCATGGGCCAT	141885	Db	142894	GGCCC--TAAAGAGGGGCCCTAACCCAGCCCGAGTGTTAAGCTTGGCAA-GGGGCAGAGAC	142950
QY	2358	TATACACTGAACATAGAGGAAGGAACACCCGTTTGTGTGTCGCCCTGCTTGAAGGAAGAA	2417	QY	3438	TTTTSTTVATAYRTACAGAAAACAAACAGRAAYAGCTCTGAGGTCTTACACAGTCCR	3497
Db	141886	TATACACTGAACATAGAGGAAGGAACACCCATTTTGTG-CCCCCTGCTTGAAGAAAGAA	141944	Db	142951	TTTTCTTTATATGTACAG-AAAAAACAGGAATAGCTCTAGAGCCATCACACAGTCTG	143009
QY	2418	TTAATCCTGAAGTCTGGGCAACAGAGGACAAATATGAGCGAGCCAAAGAAATGSCCGTCTCT	2477	QY	3498	AGGAYGAGCTTGCAACCCYRTGGCRYACCTGAATAAGGAAAYTGATGTAGTGGCAAAGGG	3557
Db	141945	TTAATCCTGAAGTCTAAGCAACAGAGGACAAATATAGACGAG-CAAGAAATGCCCATCTCT	142003	Db	143010	AGGACTAGCTTGCAACCCGTTGGCATACCTCAGTAAAGAAATTGATGTAGTACAAAGGA	143069
QY	2478	GTTCAAGTTAAACTAAGAGATTCCACTTCTCTTTCCCTACCAAGGCAAGTACCCCTCAGA	2537	QY	3558	TTGRCYTCATTTGTTTAYGGGTAGTGGGTAGTGGAGTACAGTACAGTCTGAAGCAGTTAA	3617
Db	142004	GTTCAAGTTAAACTAAGAGATTCCGCTCTTTCCCTACCAAGGCAAGTACCCCTTAGA	142063	Db	143070	TTGGCTTCATTTGTTTACAGGTAGTGGGTAGTGGAGTACAGTCTAGTGTCTGAAGTGGTTAA	143129
QY	2538	CCCAAGGCCCAACAGGATTCCAAAGATTGTTAAGGACTTAAAGCCCAAGGCTTAGTA	2597	QY	3618	AATAATCAGGGRAGAGATCTTACTGTGTGACACATCTCATKAGTGAAYRGACATCTCAC	3677
Db	142064	CCCAAGGCCCAACAGGACTCCAAAACCTGTAAAGGACCTTAAAGCCCAAGGCTTAGTA	142123	Db	143130	AATGATCAGGGAAGAGATCTTACTGTGTGACATCTCATGATGGAATGCCATCTCAC	143189
QY	2598	AAACATGCATAACTCCCTGACGATTAATCCGTTAGTGGATTGAGGAGGCACAGAAACCCAG	2657	QY	3678	TGCTAAAGGAGACTTGTGGCTGTGAGCAACACGTTTACTTAAATRTCAGGCTCTATTACT	3737
Db	142124	AAACATGCATGAGGCCCTGCAATATCTCC-----AAGTGTAGGAGTACAGAAACCCAA	142176	Db	143190	TGCTAAAGGAGACTTGTGGCTGTGAGCAACACGTTTACTTAAATRTCAGGCTCTATTACT	143249
QY	2658	TGACAGTGTGAGGGTTAGTGCAGAGATCTCAGGATTTATCAATGGAGGCCGCTTGTCTTTTA	2717	QY	3738	TGAARGCCAGTGTGCBACCTGTGCACTTGTGCAACTCTTAAACCAGYCNATTTCTTCC	3797
Db	142177	CGGACAGTGA--GGTTAGTGCAAAATCTCAGGATTTATCAAT--GAGACTGTTGTTCTCTA	142234	Db	143250	TGAAGGCCAGTGTGCAACTGCGCACTGTGTGTAACCTCTTAAACCAGTCACATTTCTTCC	143309
QY	2718	TACCAGCTGTACTAGCCCTTATACTGTGMYTTCCAAATACGAGGAAGCAGAGTGG	2777	QY	3798	AGACAACTGAAGAAAGATARAAYATACTGTCAACAAFTAAATTTCTCAAACTATGCCAC	3857
Db	142235	TACCCAGCTGAACCTAAACCTTATCTGTCTTCCAAATACGAGGAAGCAGAGTGG	142294	Db	143310	AGACAATGAAGAAAGATAGAACTACTGCCAACAGTAATGCTCAAACTATGCTCCAC	143369
QY	2778	TTTACASTCTGGACCTTMAAGATGCTTCTTGATCCCTGTATCCTGACTCTCAAT	2837	QY	3858	TCGAGGGGACCTTGTAGARGTTTCCYTTGACATCCYGACCTTCAACTTGTATCTGATG	3917
Db	142295	TATATAGTCTGGACCTTAAGATGCTTCTTCTGATCCCTGTATCTCCTGATTTCCCAA	142354	Db	143370	TCGAGGGGACCTTGTAGARGTTTCCYTTGACATCCYGACCTTCAACTTGTATCTGATG	143428
QY	2838	TTCTTTGTTTGGCTTTGAAGATCTTCAAAACCCACATCTCAACTCAGCTGGACTTTTTTA	2897	QY	3918	GAACTTCCTTTGTAGAAAAGGACCTTCAAAAAGYGGGTATGCACTGTGTCAGTGAATG	3977
Db	142355	TTCTTTGTTTGGCTTTGAAGATCTTCAAACTCAAGTCTCAACTCAGCTGGACTGTTT	142414	Db	143429	GAACTTCCTTTGTAGAAAAGGACCTTCAAAAAGYGGGTATGCACTGTGTCAGTGAATG	143488
QY	2898	CCCCAAGGGTTGAGGATAGYCCCCATCTATTGTCAGGCAATTAGCCCAAGACTTGAGY	2957	QY	3978	GAAATTTGAAGATTAATCCCTCACTCCAGGAACCTAGTGTGCTGTCAGGAGAACTAATAG	4037
Db	142415	CGCCAAAGGGTTGAGGATAGYCCCCATCTATTGTCAGGCAATTAGCCCAAGACTTGAGY	142474	Db	143489	GAAATTTGAAGATTAATCCCTCACTCCAGGAACCTAGTGTGCTGTCAGGAGAACTAATAG	143548
QY	2958	CARTYMTACTCTGGACACTCTTGTCTTCTGATGAGTGAATTAATTTTGTGCTGCTG	3017	QY	4038	CCYTCAYTKGGGCACTAGAAATTTAGGAGAAAGRAAAAGGYAAATATATATACAGACTCT	4097
Db	142475	CAATTTCTACCTAGTGGACACTCTTGTCTTCTGATGATGATTAATTTTGTAGCCACC	142534	Db	143549	CCCTCAGCTCGGGTACTAGAAATCAGGGAAGGAAAGGTAATATATATACAGACTCTA	143608
QY	3018	YRTTCAGAAACCTTGTGGCCATCAAGCCACCAAGGCTCTTMAATTTCTTCGCACTGT	3077	QY	4098	ARTATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAAGAAAGGAAATTCCTAA	4157
Db	142535	TGTTTCAGAAACCTTGTGGCCATCAAGCCACCAAGGCTCTTMAATTTCTTCGCACTGT	142594	Db	143609	AGTGTGCTTACCTAGTCTCTCCATGCCATGCCATGCAGCAATATGGAGAGAAAGGAAATTCCTAA	143668
QY	3078	GGCTACAWGGTTTCCAAACSAARGCTCACTCTGCTCAAGCAGGTTAAATCTTAGGR	3137	QY	4158	CTTTCAGGGAACACCTATACAGATCAGGAAGCCATTAGGARATTTATTATYTGCGGTAC	4217
Db	142595	GGCTACAWGGTTTCCAAACSAARGCTCACTCTGCTCAAGCAGGTTAAATCTTAGGG	142654	Db	143669	CTTTCAGGGAACACCTATACAGATCAGGAAGCCATTAGGAAATTTATTATYTGCGGTAC	143728
QY	3138	CTAARATTTACAAAGGACCCAGGCCCTCAGTGTAGGAAYRATCCAGCCTATATCTGGCT	3197	QY	4218	AGAAAACCTTARAGAGGTGGAGTCTTACACTGCGGGGTCTCATCANAAAGGAAAGGAAGG	4277

```
Db 143729 AGAAACCTARAGAGGTGGCAGTTTTCACCTGCCGGGTTCATCAGAAAGGAAAGGAAGGG 143788
Qy 4278 AATASAGGAAYTGCACGACAKATATTTGAAGMAAAGAGCTGCAAGCAGCAGCCCTC 4337
Db 143789 AATACAAAGGAGCCACCAAGTTGATTTGAAGTCAAAAGAGCCACAAGGCTGGACCCCTC 143848
Qy 4338 CATTAGAAATGCTTATTAACTTCCCTTAGTATAGGTTAATCCCTTCCGGGAACCAAGC 4397
Db 143849 CATTAGAAATGCTTATTAGGAGGCCCTTAGTATGGGGTAATCCCTTCCGGGAAGCCCAAGC 143908
Qy 4398 CCCAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGG---CAGTTTTCCTCCCT 4454
Db 143909 CCCAGTACTCAGCAGGAGAAATAGATAGGGAACCTCATGAGGACATCTTCCCTCCCT 143968
Qy 4455 CGGACGGTTAGCCACTGAAGAAGGGGAAATATCTTTTGCCTGCAATCATCAATGAAAT 4514
Db 143969 CCAGATGGCTAGCCACCAATAAAGGAAATACTTTTGCCTGCGAGCTAACCAATAGAAAT 144028
Qy 4515 TACTTAAACCCCTTCATCAACCTTTTCACTTAGGCTCGATAGCAGCCATCABATGGCCA 4574
Db 144029 TACTTAAACCCCTTCATCAACCTTCACCTTAGGCTATGATAGCCCATGAGATGGCCA 144088
Qy 4575 AATCATTTATTACTGACACAGGCTTTTCAAACTATCAAGCARATAKTCAAGGCGCTGTG 4634
Db 144089 AATTATTATTACTGACACAGGCTTTTCAAACTATCAAGCARATAKTCAAGGCGCTGTG 144148
Qy 4635 AATGTGCCARAAATATFCCCTGCTVATCGCAAGCTCCTTCA 4681
Db 144149 AAGTCTGCCAAAGAAATATATCCCTGCTGCTGACGAGGCCATACATTCA 144195

RESULT 12
US-11-121-086-78
; Sequence 78, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78
; LENGTH: 189993
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-78

Query Match 41.5%; Score 3105.4; DB 15; Length 189993;
Best Local Similarity 86.7%; Pred.No. 0;
Matches 3580; Conservative 146; Mismatches 313; Indels 88; Gaps 29;

Qy 606 GATGGAAAGCTTCCCGCAAGACAAAGCGCCCTTAAGACGTATTTCTGGARAAATGGGA 665
Db 9479 GATGGAAATGTTTCCCCCAAGCAGAAATGCCCTTAAGATGTATTTCTGGAGAAATGGGA 9538
Qy 666 MCAATTTGACCTCAGACACTAAGAAAGAAACAGCTTATTTCTTCGAGTGCCGCTG 725
Db 9539 CCAATCTGACCATCAGACACTAAGAAAGAAATGACTTATATTTCTTCGAGTACCACTG 9598
Qy 726 GC-----ACTCCTGAGGGAAGTATAAATTTATAA 753
Db 9599 GCACAATATCTTCTTCAAGGGGCGAAGAACCTTGGCCCTCTCGAGGGAAGTATAAATTTATAA 9658
Qy 754 CACCATCTTACAGCTAGACVTCTTTTGTAGAA---AAGGCAATGGAGTGAAGTGCATA 810
Db 9659 CACCATCTTACAGCTAGACCTCTTTTGTAGAAAGGAAAGCAATGGAGTGAAGTGCATA 9718
Qy 811 AGTACAAACCTTTCTTTTCATTAAGAGACAACTCACAATTTATGTAAAAAGTGTGTTATG 870
```

```
Db 9719 TGTACAAACCTTTCTTTTTCATTAAGAGATACTCCCAATTTATGTAAAAAGTGTGATTTATG 9778
Qy 871 CCCTACAGGAAGCTTTCCAGAGTCTACTCCCTATCCAGCA--TCCCGGACTCTCTCCCCC 928
Db 9779 CCCTACAGGAAGCCCTCAGAGTCTACTCCCGACCCCGCAGCAGACCCCAACTCTCTCTCC 9838
Qy 929 AMYTAATAAGGACCCCTTTCAACCCAAATGGTCCAAAGAGGATAGACAAAAGGGTAAA 988
Db 9839 AACTAATAAGGACCCCTTTCAACCCAAATGGTCCAAAGAGGATAGACAAAAGGGTAAA 9898
Qy 989 CAGTGAACCAAGAGTCCCAATATTTCCCAATTTATGACCCCTCCCAAGCAGTGGGAGAA 1048
Db 9899 CAATGAACCAAGAGTCCCAATATTTACAGATTTATACGATTTATCTCGCTCC---AAGCAGTGGGAG 9954
Qy 1049 GAGAAATTCGCGCCAGCAGAGTGCATGTGCYTTTCTCCAGACTTAAAGCAATAA 1108
Db 9955 GAGAAATTCGCGCCAGCAGAGTGCATGTGCYTTTCTCTCTCAGATTTAAAGCAATAA 10014
Qy 1109 AAACAGACTTAGGTAAATTTCTCAGATAAATCTGATGGCTATATTTGRTGTTTTTCAAGGGT 1168
Db 10015 AATAGACCTTAGGTAAATTTCTCAGATAAATCTGATGGCTATATTTGATGTTTTTCAAGGGT 10074
Qy 1169 TAGGACAATTTCTTTGATCTGACATGAGAGATATATATGTCATGCTTAATCAGACACTA 1228
Db 10075 TAGGACAATTTCTTTGATCTGACATGAGAGATATA-ATGTTACTGCTTAATCAGACACTA 10133
Qy 1229 ACCCCAAATGAGAGAAGTGCACCAATAACTGACGCTGAGAGTTTGGCGACTCTCGTAT 1288
Db 10134 ACCCCAAATGAGAGAAGTGCACCAATAAGCTGCTGATGAGAGTTTGGCGAACTCTCGTAT 10193
Qy 1289 CTCAGTCAGGTCAATGGATANGATGACAAACAGAGAAAGAAATGATTTCCCAACAGGC 1348
Db 10194 CTCAGTCAGGTCAATG--ATAGGATGACAAACAG-ATGAAAGAGAAATGATTTCCCAACAGGC 10250
Qy 1349 CAGCARGCAGTTCCCGAGTCTASACCTCATTTGGGGGACAGCAAAATCAGTAAATGGGAGA 1408
Db 10251 CAGCARGCAGTTCCCGAGTCTAGACCTCAT--AGGACACAGAA---TCAGAACTTGGAGA 10306
Qy 1409 TTGGTGCTCAGACACTTTCTTACTTGTGCTGCTASAGGACTTAAGGAAACTAGGAGAA 1468
Db 10307 TTGGTGCCACAGACATTTCTTAACTTGTGCTGCTAGAGGACTTAAGGAAACTAGGAG-A 10365
Qy 1469 ARTTAYGAATTACTCAATGATGTCCACCATTAACACAGGGGAGGAAAGAAATCTCTACT 1528
Db 10366 AGCCCATGAATTTATCAATGATGTCCCTTAAACACA-GGAAAGGAGAAATCTCTACT 10424
Qy 1529 GCCTTTCTGGAGAGACTAAGGAGGCAATTTAGAGAACGCTGCTCTCTGTCACCTGACTCT 1588
Db 10425 GCCTTTCTGGAGAGACTAAGGAGGCAATTTAGAGAACGCTAATCCTCCTGTCACTGACTCT 10484
Qy 1589 TCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCAGTCAGTCAGCTGAGACATTT-AG 1647
Db 10485 ATTAAAGGCCAACTAATCTTAAAGGATTAAGTTTATCAGTCAGTCAGCTGAGAGATTAAG 10544
Qy 1648 AAAAACTTTCAAAAGTCTGCCGTAGCGCCGAGCAAAAATTTAGAAAACCTTATTGAAGCTTG 1707
Db 10545 AAAAACTTTCAAAAGTATGCTTAGGCCCAGAGCAAAAATTTAGAAAACCTTACTGAAGCTTG 10604
Qy 1708 GCAACTTCGGTTTTTTTATTAATAGAGATCAGAGGAGCAGCGGGAACAGGACAAACCGGAT 1767
Db 10605 GCAACTTCAGTTTTTTTATAATAGAGATCAGGAAGAGCAGG-GGAATGGGACAAATGGGAT 10663
Qy 1768 -----TAAAAAAGGCCACCGCTTAGTATCATGCCCTCAGGCAAGTGGACTTTGGAG 1820
Db 10664 AAAAAAAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 10723
Qy 1821 GCTCTGGAAAAGGAAAAGCTGGGCAAAATTTGAATGCTTAATAGGGCTTCTCTCCAGTGGC 1880
Db 10724 GCTCCAGAAAAGGAAAAGCTGAGCAAAATTTGAATGCTTAACAGGGCTTCTCTTAGTGTG 10783
Qy 1881 GTCTACAAGGACACTTTTAAAAAAGATTTGTTCAAGTAGAAGTAAAGCGCCCTTCTGTCAT 1940
```

Db	10784	GTCTACAAGGACACCTTTAAATAAAGATTGTTCACAGTAGAATAAACAAGCTGCCCCCTTCTCCAT	10843
Qy	1941	GCCCTTATTCAAGGGAACTACTGGAAGGCCCACTGCCCAAGGGGACAAGGCTCTTTTG	2000
Db	10844	GCCCTTATGTCAAGGGAATCACTGGAAGGCCCACTGCCCAAGGGAGTGAAGGCTCTCTG	10903
Qy	2001	AGTCAGAAGCCACTAAACAGATGATCCAGCAGCAGACTGAGGGTGCCTGGGGCAAGGC	2060
Db	10904	AGTCAGAAGCCACTAAACAGATGATCCAGCAGCAGACTGAGGGTGCCTGGGGCAAGGC	10963
Qy	2061	CATCCATGCCATCAACCTTCACAGAGCCCTGGGTATGCTTGACCAATTGAGGGCCAGGAAG	2120
Db	10964	CAGGCCATGCCATCAACCTTCACAGAGCCCTGGGTATGCTTGACCAATTGAGGGCCAGGAAG	11023
Qy	2121	GT---TGTCCTCGACACTGTGGGTCTTCTTAGTCTTACTCTCTGTGTCGGGACAAC	2177
Db	11024	TTCACTGTCTCTTGACACTGTGTGGCCCTTCTCAGTCTTACTCTCTCTGTCTGTGACAAC	11083
Qy	2178	TGTCCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGCAGTCACTAGATACT	2237
Db	11084	TGTCCTTCAGATCTGTCACTATCC---GAGGGGTCTTAGGACAGCCAGTCACTAGATAC-	11139
Qy	2238	TTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTTTTCATGCTTTTCTAAT	2297
Db	11140	TTCTCCAGCCACTGAGTTGTG-ACTGGGGAACCTTTACTCTTTTTCATGCTTTTCTAAT	11198
Qy	2298	TATGCTTGAAGCCCACTACCTTGTGTAGGGAGAGACATTTAGCAAAAGCAGGGGCCAT	2357
Db	11199	TATGGTGAAGGCCCACTCCCTTGTGTAGGGAGAGACATTTAGCAAAAGCAGTGGGCCAT	11258
Qy	2358	TATACACCTGAACATAGGAGAGAACCCGTTTGTCTGTCNCCCTGCTTGAAGGAAGAA	2417
Db	11259	TATACACCTGAACATAGGAGAGAACCCATTTGTGTG-CCCTGCTTGAAGGAAGAA	11317
Qy	2418	TTAATCCTCAAGTCTGGGCAACAGAGAGACAATATGAGAGGCCCAAGAAATGCCGTCT	2477
Db	11318	TTAATCCTCAAGTCTAAGCAACAGAGAGACAATATAGACGAG-CAAGAATGCCATCCT	11376
Qy	2478	GTTCAGTTAAACTAAGGATTCCTCAAGATTTGTGTAGGACTTAAAGCCCAAGGCTTAGTA	2537
Db	11377	GTTCAGTTAAACTAAGGATTCCTCAAGATTTGTGTAGGACTTAAAGCCCAAGGCTTAGA	11436
Qy	2538	CCCAAGGCCCAAGAGATTCCTCAAGATTTGTGTAGGACTTAAAGCCCAAGGCTTAGTA	2597
Db	11437	CCCAAGGCCCAAGAGATTCCTCAAGATTTGTGTAGGACTTAAAGCCCAAGGCTTAGTA	11496
Qy	2598	AAACCATGCATAAATCCCTGCAGTAATTCCTGATGGAATTGAGGAGGCCAGAAACCCAG	2657
Db	11497	AAACCATGCAGTAGCCCTGCAATCTCC-----AACTGTAGGAGTACAGAAACCCAA	11549
Qy	2658	TGACAGTGGAGGTTAGTGCAAGATCTCAGGATTAATAGGAGGCCGTTGTCTTTTA	2717
Db	11550	CGACAGTGGGA-GGTTAGTGCAAAATCTCAGGATTAATCAAT-GAGACTGTGTCTCTTA	11607
Qy	2718	TACCCAGCTACCTAGCCCTTATCTGTGMYTTCCCAATACCAGAGAACGAGTGCG	2777
Db	11608	TACCCAGCTACCTAGCCCTTATCTGTGMYTTCCCAATACCAGAGAACGAGTGCG	11667
Qy	2778	TTTACASTCTGGACCTTMAAGATGCTCTTCTGCATCCCTGTATCATCTCACTCAAA	2837
Db	11668	TATATAGTCTGGACCTTAAGGATGCTTCTTCTGCATCCCTGTATCATCTGATTTCCAA	11727
Qy	2838	TTCTGTGTTTGCCTTTGAAGATCTTCAAAACCCARCATCTCAACTCACTGGATTTTAA	2897
Db	11728	TTCTGTGTTTGCCTTTGAAGATCTTCAAAACCCARCATCTCAACTCACTGGATTTTAA	11787
Qy	2898	CCCCAGGCTTCAGGATAGYCCCAATCTATTGGCCAGGCATTTAGCCCAAGCTTGAGY	2957
Db	11788	CGCCAAAGGTTACAGGATAGCCCCCATCTATTGGCCAGGCATTTAGCCCAAGATTTGAGC	11847
Qy	2958	CARTYMTCACTACCTGACACTCTTGTCTTTCRGTAKGTGGATGATTTACTTTTTCGCGC	3017
Db	11848	CAATTCTCATACCTGGACACTCTTGTCTTTCRGTAKGTGGATGATTTACTTTTTCGCGC	11907
Qy	3018	YRTTCAGAAACCTTGTGCCATCAAGCCACCACCAAGCRCTCTTMAATTTTCTCGCYACCTGT	3077
Db	11908	TGTCAGAAACCTTGTGCCATCAAGCCACCACCAAGGTCTTTAAATTTCTCACCACTGT	11967
Qy	3078	GGCTACAGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGCAGTTTAAATPACTTAGGR	3137
Db	11968	GGCTACAGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGCAGTTTAAATPACTTAGGG	12027
Qy	3138	CTAATATTATCAAAAGGACCAARGCCCTCAGTGAAGAAAYATCCAGCCTTACTTGGCT	3197
Db	12028	CTAATATTATCAAAAGGCTTGAAGGCCCTCAGTGAAGAAACATCCAGCCTTACTTGGCT	12087
Qy	3198	TATCCTCATCYCAAAACCCCTAAAGCAACTAAGRBRRTTCTTGGCCTTAAYAGGTTTTCG	3257
Db	12088	TATCCTCTTCCAAACCCCTAAAGCAACTAAGAGGGTTCCTTGGCATACAGGTTTCTAC	12147
Qy	3258	CGAATATGGATTCCCCAGGTGTGGCRAAATAGCCAGGYCATTTAMATACASTAATTAAGGA	3317
Db	12148	CGAATGTGGATT-CCAGGTACAGCAAAATAGGCAGATCATTTATATACATAATTAAGGA	12206
Qy	3318	AACTCAGAAAGCCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAGTGCGCTTCCA	3377
Db	12207	AACTCAGAAAGCCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAGCAGCTTCCA	12266
Qy	3378	GGCCCTTAAAGAGGCCCTTAAACCAAGVCCAGTGTAAAGVTTGCCAACRCGGCAAGAC	3437
Db	12267	GGCCC--TAAAGAGGCCCTTAAACCAAGCCCGGTGTAAAGTGTGCCAA-GGGGCAAGAC	12323
Qy	3438	TTTTSTTTATATYRTTCACAGAAATAACAGRAAYAGCTCTTRGGAGTCTTTTACACAGTCCR	3497
Db	12324	TTTTCTTTATGTGCACAG-AAAAAACAGGAATAGCTCTAGGAGCCATCACAACAGTCTG	12382
Qy	3498	AGGAYAGCTTGCACACCTGTCGRYACCTGASTAAGAAAYTGATGTAGTGCGAAAGGG	3557
Db	12383	AGGACTAGCTTGCACACCTGTCGATACCTGAGTAAAGAAATGTAGTAGTACGAAGGA	12442
Qy	3558	TGRCYCTCATTTTAYGGGTAGTGTGCGTAGTACAGTGTAGTATCTGAAGACAGTTAA	3617
Db	12443	TGGCTTTTATTTTACAGGTAGTGTGCGTAGTACAGTGTAGTGTCTGAAAGTGTTAA	12502
Qy	3618	AATAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAGTGAAYRGATACCTAC	3677
Db	12503	AATGATACAGGAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCCATCTCAC	12562
Qy	3678	TGCTAAAGAGACTTGTGGCTGTGCAGCAACVGTGTTTACTTTAAATRTCAGGCTCTATTCT	3737
Db	12563	TGCTAAAGAGACTTGTGGCTGTGCAGCAACCGTTTACTTTAAATATCAGGCTCTATTCT	12622
Qy	3738	TGAAGGCCAGTGTGCRACCTGTGCATCTGTGCAACTCTTAAACCCAGYCNCAATTTCTCC	3797
Db	12623	TGAAGGCCAGTGTGCRACCTGTGCATCTGTGCAACTCTTAAACCCAGTCAATTTCTTC	12682
Qy	3798	AGACATGAAGAAAGATARAAYATAAAGTGTCAACAARTAAATTTCTCAAAACCTATGCCAC	3857
Db	12683	AGACATGAAGAAAGATARAAYATAAAGTGTCAACAARTAAATTTCTCAAAACCTATGCCCG	12742
Qy	3858	TCGAGGGACCTTGTAGAGTTCCVTTGACTGATCCYGACCTTCAACTTGTGATPACTGATG	3917
Db	12743	TCGAGGGACCTTGTAGAGTTCCVTTGACTGATCCYGACCTTCAACTTGTGATPACTGATG	12801
Qy	3918	GAAGTTCTTTGTAGAAAGGACTTCGAAAGYGGGTATGCAGTGTGCAGTGAATAG	3977
Db	12802	GAAGTTCTTTGTAGAAAGGACTTCGAAAGYGGGTATGCAGTGTGCAGTGAATAG	12861
Qy	3978	GAATATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTTRGCGAGAACTAATAG	4037
Db	12862	GAATATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTTRGCGAGAACTAATAG	12921
Qy	4038	CCYTCATYKGGGCACTAGAAATTTAGGAGAAAGGAAAGGGYAAATATATATACAGACTCTR	4097
Db	12922	CCCTCACTCGGGTACTAGAAATTTAGGAGAAAGGAAAGGGYAAATATATATATACAGACTCTA	12981



```
QY 4098 ARTATGCTYACCTAGTCNTCCATGCGCCCATATGSRAGAAAGGATTCCTAA 4157
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
12982 AGTGTGCTTACCTAGTCCTCATGCCCATGCGCAGCATATGAGGAAAGGAATTCCTAA 13041
QY 4158 CTTTCYAGRGAACACCTTATCAMACATCAGGAAGCCATTAGGARATTTATTTGCGWGTAC 4217
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13042 CTTCCGAGGGAACACCTTATCAACATCAGGAAGCCATTAGGAAATTTATTTGGCTGTAC 13101
QY 4218 AGAAACCTAAGAGGTGMAAGTCTTACATCTGCGGGTCAATCAAAAGGAAAGGAAGGG 4277
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13102 AGAAACCTAAGAGGTGCAAGTTTACATCTCCGGGGTCTATCAAAAGGAAAGGAAGGG 13161
QY 4278 AATATASAGGAAYTGCAAGCAKATATGAAGCMAAAAGAGCTGCAAGGACGAGCCCTC 4337
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13162 AATACAAGGAGGCCCAAGTTGATATGAAGTCAAAAGGCCAAGGCTGAGCCCTC 13221
QY 4338 CATTAGAAATGCTTATTTAAATCTTCCCTTAGTATAGGGTAAATCCCTTCCGGGAAACCAAGC 4397
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13222 CATTAGAAATGCTTATAGGAGGACCCCTAGTATGGGGTAAATCCCTTCCGGGAGCCDAGC 13281
QY 4398 CCCAGTACTCAGCAGGAGAAACAGATGGGGAACCTCACGAGG---CAGTTTTCTCCCT 4454
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13282 CCCAGTACTCAGCAGGAGAAATAGAATAGGGAACCTTCATGAGGACATATCTCCCTCCCT 13341
QY 4455 CGGAGCGTTAGCCACTGAAGAGGGAAGAAATATCTTTTGCCTGCAACTATCCAATGGAAT 4514
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13342 CCAGATGGCTAGCCACCATAAAGGAAATAATCTTTTGCCTGAGCTTAACCAATAGAAAT 13401
QY 4515 TACTTAAACCCCTTCATCAAACTTTTCACCTTAGGCATCGATAGCACCATCARATGGCCA 4574
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13402 TACTTAAACCCCTTCATCAAACTTTTCACCTTAGGCATGTATAGACCCATGATGGCCA 13461
QY 4575 AATCATTTTACTGGACGAGGCTTTTCAAACTATCAAGCARATATACTCAGGCGCTGTG 4634
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13462 AATTATTATTACTGGACGAGGCTTTTCAAACTATCAAGCAGATAGTCAGGCGCTGTG 13521
QY 4635 AATGTGCCARAAATATCCCTGCTCTATCGCAAGCTCCTTCA 4681
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13522 AGTCTGCCAAAGAAATAATCCCTGCTGCACTGCGAGGCCATACATTTTCA 13568
```

## RESULT 13

```
US-09-997-722-148
; Sequence 148, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 148
; LENGTH: 22436
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-997-722-148
```

```
Query Match 40.7%; Score 3049.8; DB 3; Length 22436;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 3556; Conservative 142; Mismatches 320; Indels 93; Gaps 31;
QY 606 GATGGGAAAGCTTCCCGGAAGACAAACCGCCCTTAAGACGTATTCTGGARAAATGGGA 665
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1581 GGTAGCAACATCTCCCTCAAGCAAAACACCCCGAAGATGTATTTCTGGAGATCGGA 1640
QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGCACTGCCGCTG 725
```

```
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1641 CCAATTTGACCTCAGACGCTGAGAAAGAAATGACTTATATTTCTTCGAGTATTGCGTG 1700
QY GC-----ACTCCTGAGGAAAGTATAAATTATA 753
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1701 GCCACAATATCCTCTTCATGCGGGGAGAAACGTGGCCTCCTGAGGGAAGTATNAATTATGA 1760
QY 754 CACATCTTACAGCTAGACVCTTTTGTGAGAA---AAGGCAATGGAGTGAAGTGCATTA 810
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1761 CACATCTTACAGCTAGACCTCTTTTGTAGAAAGAGGCAATGGAGTGAAGGCCCATTA 1820
QY 811 AGTACAAACCTTCTTTTCAATTAAGACAACTCACAATTTATGTAAAGAGTGTGATTTATG 870
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1821 TGTACAAATTTCTTTTCAATTAAGACAACTCGCTATTATGTAAAGAGTGTGATTTATC 1880
QY 871 CCTACAGGAAGCCTTTCAGAGTCTACCTCCTATCCAGCAT---CCCCGACTCTCTCCCC 928
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1881 CCCTACAAGAGCCCTCAGAGTGTACTTCTTACCCAGCATCCCCCTACTCTCTCCCC 1940
QY 929 AMYTAATAAGGACCCCTTCAACCCAAATGTCTCAAAGGAGATAGACAAAAGGTTAAA 988
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1941 AACTAATAAGGACCCCTTCAACCCAAATAGTCCAAAGGAGATAGACAAAAGGTTAAA 2000
QY 989 CAGTGAACCAAGAGTGCCTTCCCAATTTATGACCCCTCCCAAGCACTGGAGGAA 1048
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2001 CAATGAATCAAGAGTGCCTTCCCAATTTATGCGAATTTATGCCACT-CCAAACAGTGGAGGAG 2059
QY 1049 GAGAAATCGGCCCAGCAGAGTGCATGTCTTTTCTCCAGACTTAAAGCAATAA 1108
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2060 GAGAAATCGGCCCAGCAGAGTGCATGTACATTTCTCTCAGCTTAAAGCAATAATG 2119
QY 1109 AAACAGACTTAGGTAATTTCTCAGATAAATCTGATGGCTATATTTGTTTCAAGGGT 1168
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2120 AAATAGACCTAGGTAATTTCTCAGCTAAACCTGATGGCTATATTTGATTTTACAAGGGT 2179
QY 1169 TAGGACAATTTTGTATCTGACATGAGAGATATATGTCATCTGCTAAATCAGACACTA 1228
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2180 TAGGACAATCTTTGATCTGACATGGAGAGATATA-ATGTTACTGTTAAATCAGACACTA 2238
QY 1229 ACCCAATAGAGAAAGTGCACCAATACTGACGCTGAGGTGTTGCGCATCTCTGTAT 1288
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2239 ACCCAATAGAGAAAGTGCAGCAATCTGACGCCAGAGGTTGCGCATCTCTGTAT 2298
QY 1289 CTCAGTCAGTCAATGGATANGGATGACAAAGAGGAAAGAAATGATTTCCCAAGGC 1348
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2299 CTCAGTCAGTCCAT--CATAGGATGAAACAG-AGGAAAGAGAACGATTTCCCAAGGC 2355
QY 1349 CAGCAGGAGTTCAGTCTASACCTCATTTGGGACACAGAAATCAGTAAATGGGAGA 1408
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2356 CAGCAGGAGTTCAGTCTAGACCTCATAGGG---ACACAGAAATCAGAAAT-GGAGA 2411
QY 1409 TTGGTCTCAGACATTTGCTTAACCTGCTGCTASAGGCTAAGGAAACTAGGAGAA 1468
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2412 ATGGTGCAGACATTTGCTTAACCTGCTGCTAGAGGACTAAGGAAACTAGGAGAA-A 2470
QY 1469 ARTTAYGAATTAATCAATGATGTCCACCAATAACACAGGGGAGGAAAGAAATCTTACT 1528
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2471 AGCCTATGAATTTATCAATGATGTCCAGTATAACA-AAGGAAAGAGGAAATCTTACT 2529
QY 1529 GCCTTTCTGGAGAGACTAAGGAGGCAATTTAGGAAAGCGTCCCTCTCTGTCACTGACTCT 1588
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2530 GCCTTTCTGGAAAGACTAAGGAGGCAATTTAGGAGGAAACCTGTCTGTCACTGACTCT 2589
QY 1589 TCTGAAGGCCAATTAATCTTAAAGCGTAAATTTATCACTCAGTCAGCTGAGCACTAG- 1647
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2590 TTTGAAGGCCAATTAATCTTAAAGGAAAGTTTATCACTCAGTCAGCTGAGCACTAGTA 2649
QY 1648 AAAAATCTTCAAAAGTCTCCCTAGGCCCGGAGCAAACTTAGAAACCTCTATTGAAGTTC 1707
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2650 AAAAATCTTCAAAAGTCAAGCTTAGGCCCTGGAGCAAACTTAGAAACCTCTATTGAAGTTC 2709
QY 1708 GCAACVTCGG-TTTTTTATAATAGAGATCAGGAGGAGCGGCGGAAACAGGACAAACGGGA 1766
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```



Db 2710 GCAACCTCGGTTTTTTTATATAGATCAGGAGGAGCGGGAATGGACAAACGGGA 2769  
Qy 1767 TTAATAAAAGGCCACCGCTTTAGTCAATGACCCCTCAGGCAAGTGAACCTTTGGAGGCTCTG 1826  
Db 2770 -TAAAAAAAGGCCACCGCTTTAGTATATAGCCCTCAGGCAAGTGAACCTTTGGAGGCTCTG 2828  
Qy 1827 GAAAGGGGAAGCTGGGCAATTCGATGACCTAATAGGCTTGCCTTCAGTGGCGTCTAC 1886  
Db 2829 GAAAGGGGAATTTGCTGAGCAAAATGGGTGGCTTAAATAGGCTTGCCTTCAGTGGCGTCTAC 2888  
Qy 1887 AAGGACACTTTAAAAAGATTCGCAAGTAGAAGTAAGCGGCCCTCTCGTCCATGCCCT 1946  
Db 2889 AAGGACATTTAAAAAGATTCGCAAGTAGAAGTAAGCGGCCCTCTCGTCCATGCCCT 2948  
Qy 1947 TATTTCAAGGGAATCACTGGAAGGCCACTGCCCCAGGGGAACAAAGGCTTTTGTAGTCAG 2006  
Db 2949 TATATCAAGGGAATCACTGGAAGGCCAGGGGCCCGGGGGATGAAGGTCCTCTGAGTCAG 3008  
Qy 2007 AAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGCTGGGCGGCAAGGCCATCCC 2066  
Db 3009 AAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGCTGGGCGGCAAGGCCATCCC 3068  
Qy 2067 ATGCCACTCACTCAGAGCCCTGGGTATGCTTGACCATTCAGGCGCCAGGAAAGT--T 2123  
Db 3069 ATGCCACTCACTCAGAGCCCTGGGTATGCTTGACCATTCAGGCGCCAGGAAAGT--T 3128  
Qy 2124 GTCTCTGGACACTGGTGGGCTTTCTTTAGTCTTACTTTCTGTCGCCGACAACTGTCT 2183  
Db 3129 GTCTCTGGACACTGGCAGCGCTTCTCAGTCTTACTTCTGCTCTGGACAACTGTCT 3188  
Qy 2184 CCAGATCTGCTACTATTCTGAGGGGTCNTAAGACGGGCACTACTAGATCTTTTTC 2243  
Db 3189 CCAGATCTGCTACTATCC---GAGGGGCTCTAGGACAGCAGTCACTAGATAC-TTCTCC 3244  
Qy 2244 CAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTTCACATGCTTTTCTAATTATGCT 2303  
Db 3245 CAGGCATTAAGTTATG-ACTGGGAACCTTACTCATTTTCATGCTTTTCTAATTATGCC 3303  
Qy 2304 TGAAGCCCCACTACCTGTTTAGGGAGACATTTCTAGCAAAAGCAGGGGCCATTATACA 2363  
Db 3304 TGAAGCCCCACTCCCTTGTAGTGAGAGAAATTTCTAGCAAAAGCAGGGGCCATTACACA 3363  
Qy 2364 CCGTAACATAGAGAAGGAACACCGCTTTGTTGTCNCCCTGCTTGAAGGAAGGAATTAATC 2423  
Db 3364 CCGTAACATAGAGAAGGAACACCGCTTTGTTGT-CCCCCTGCTTGAAGGAAGGAATTAATC 3422  
Qy 2424 CTGAGCTCTGGCAACAGAAAGACAAATATGGAGGACCAAGAAATGCCCGTCTCTGTTCAA 2483  
Db 3423 CTGAGCTCTGGCAACAGAAAGACAAATATGGAGGACCAAGAAATGCCCGTCTCTGTTCAA 3480  
Qy 2484 GTTAAACTAAGGATTTCCACTTCTCTTCCCTACCAAGGCACTACCCCTCAGACCCCAAG 2543  
Db 3481 GTTAAACTAAGGATTTCCACTTCTCTTCCCTACCAAGGCACTACCCCTCAGACCCCAAG 3539  
Qy 2544 GCCCAACAAGGATTCGCAAGATTTAAGGACTTAAAGCCCAAGCCCTAGTAAACCA 2603  
Db 3540 GCCCAACAAGGATTCGCAAGATTTAAGGACCGGAAAGCCCAAGCCCTAGTAAACCA 3599  
Qy 2604 TGCATAACTCCCTGCAAGTAAATCCGTAGTGGATTGAGGAGGACAGAAACCCAGTGGACA 2663  
Db 3600 TGCATAAGCCCTTGCATATTTCC-----AATTTTAGAGATACAGAAACCCAGTGGACA 3652  
Qy 2664 GTGGAGGGTTAGTGCAGATCTCAGGATTAATAGGAGCGGTTGCTCTTTTATACCCA 2723  
Db 3653 GTGGA-GGTTAGTGCAGATCTCAGGATTAATGAT-TAGGCTGTTGTTCTCTATATCCA 3710  
Qy 2724 GCTGTACCTAGCCCTTATCTGTGTTTCCCAATACACAGGAAGCAGAGTGGTTTACA 2783  
Db 3711 GCTGTACCTAACCCTTATACCTGCTTCCCAATATACAGGAAGCAGAGTGGTTTACA 3770  
Qy 2784 STCTCGACCTTMAAGATGCTTCTTCTGATCCCTGTATCATCTCAGCTCTCAATCTTGG 2843  
Db 3771 GTCTGGACCTTAAGGATGCTTTTCTGATCCCTTGTATCATTTCTGACTCTCAATTTTATA 3830

Qy 2844 TTTGCTTTTGAAGATACCTTCAAAACCCARCATCTCAACTCACTGACCTRTTTTACCCCAA 2903  
Db 3831 TTTGCTTTTGAAGATCCTTTGCAACCCAAATGCTCAACTCACTGACCTGTTTTTACCCCAA 3890  
Qy 2904 GGGTTACGGGATAGYCCCCCATCTATTGGCCAGGCAATTAGCCCAAGACTTGAGYCARVM 2963  
Db 3891 GGGTTACAGATAGCCCCCATCTATTGGCCAGGCAATTAGCCCAAGACTTGAGCCAAATTC 3950  
Qy 2964 TCATACCTTGGACACTTTTGTCTTCRGTAKGTGGATGATTTACTTTTTRGTCGCCVTTTCA 3023  
Db 3951 TCATACCTTGGACACTTTTGTCTTCRGTAKGTGGATGATTTACTTTTGTAGCCCAATTTCA 4010  
Qy 3024 GAAACCTTGTGGCATCAAGCCCAACCAAGRCCTCTTMAATTTCTCGCYACCTGTGGCTAC 3083  
Db 4011 GAAGCCTTGTGGCATCAAGCCCAACCAAGCACTCTTAAATTTCTCGCTACCTGTGGCTAC 4070  
Qy 3084 AWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGCAGGTTAAATATCTTAGGRCTAARA 3143  
Db 4071 AA-GATTCCAAACCAAGGCTCAGCTCTGCTCACAGAGTTTAAATATCTTAGGGCTAARA 4129  
Qy 3144 TTATCCAAAGGCAACARGGCCCTCAGTGAGGAAYATCCAGCCTATATCTGGCTTATCCT 3203  
Db 4130 TTATCCAAAGGTACCCAGGGCCCTCAGTGGGGAACATATGCACTATATCTGGCTTGTCT 4189  
Qy 3204 CATCYCAAAACCTTAAAGCAACTAAGRRRTTCTTGGCRTAAYAGGYTTCTGCCCAWA 3263  
Db 4190 CATCCAAACCCCT-AAGCAACTAAGAGGGTTCTTGGCATAACAGGTTTCTGCCAAATA 4248  
Qy 3264 TGGATTTCCCAAGTWTGCRAAATAGCCAGGYCATTAATATACATAATTAAGCAACTCA 3323  
Db 4249 TGGATT-CCAGGATATGGCAAAATATAGCCAGATCATTTATATACATTAATTAAGCAACTCA 4307  
Qy 3324 GAAAGCCCAATACCCATTTTARTAAAGATGGAYAMCTGAAGYMRAGTGGCTTTCCAGGCC 3383  
Db 4308 GAAAGCCCAATACCCATTTTARTAAAGGTGGACACCTGAAGCACAAGCAGCTTTCCAGGCTCT 4367  
Qy 3384 TAAAGAGGCTTAAACCAAGYCCAGTGTAAAGTTTGGCCAAACGGGCAAGACTTTTCT 3443  
Db 4368 AAAGAA--GGCTCTAACCCAAAGTTCCAGTGTTCACGTTTGGCCAAAGGGCAAGATTTTCT 4425  
Qy 3444 TVATAYRTACAGAAAAAACAAGRAAYAGCTCTCGAGTCTTACACAGRTCCRAGGAY 3503  
Db 4426 TTAATGTCAACAGAAAAAACAAGAAATAGCTCTAGAGTCTTACACAGTCCGAGGAT 4485  
Qy 3504 GAGCTTGCAACCCYRTGGCRYACCTGAATAAGGAAAYTGATGTAGTGGCAAGGGTTGRCY 3563  
Db 4486 GAACTTGCCAACTGTGGCATACCTGAGTAAGGAAATTGATGCAGTGGCAAGGGTTGGCC 4545  
Qy 3564 TCATTTGTTAYGGGTAGTGGCGGCTAGTACAGTGTAGTATCTGAAGCAGTTAAATAAT 3623  
Db 4546 TCATTTGTTATGGGTAGTGGCGGCTAGCAGTGTAGTATCTGAAGCAGTTAAATAAT 4605  
Qy 3624 ACAGGAGAGATCTTACTGTGTGACATCTCATGAKGTGAAYRGATACACTACTGCTAA 3683  
Db 4606 ACATGGAAGAGATCTTACTGTGTGACATCTCATGATGTGAACGGCATACTACTGCTAA 4665  
Qy 3684 AGGAGACTTGTGGCTGTACAGCAACYGTTTACTTAAATRTCAGGCTCTAATTAAGTCAARG 3743  
Db 4666 AGGAGACTTGTGGCTGTACAGCGATCGTTTACTTAAACATCAGGCTCTAATTAAGTCAARG 4725  
Qy 3744 GCCAGTGTGACACTGTGCACTGTGCAACTCTTAACCCAGYCNACATTTCTTCAGACAA 3803  
Db 4726 GCGTGTGTGCAACTGCGCGCTTTGCAACTCTTAAACCCAGGCAATTTCTTCAGACAA 4785  
Qy 3804 TGAAGAAAGATARAAYATAACTGTCAACAARTAAATTTCTCAAACTATGCGCACTCGAGG 3863  
Db 4786 TGAAGAAAGATARAACATAACTGTCAACAAGTAATTTGCTCAAACTAATCCCGCTCGAGG 4845  
Qy 3864 GGACCTTTTAGAGGTTCCYTTGCACTGATCCYGAACCTTCAACTTGTATATCTGATGGAAGTT 3923  
Db 4846 GGACCTTTTAGAGGTTCCCTTGCATGATTCGACCTT-AACTTGTATATCTGATGGAAGTT 4904



QY	1588	TTCTGAAGCCCACTAACTCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAG	1647	QY	2662	CAGTGGAGGGTTAGTGCAGAGATCTCAGGATTTATCAATGGAGCCCGTGTCTCTTTTATACC	2721
Db	48728	TATTGAAGCCCACTAACTCTTAAAGAAAGTTTATCACTCAGTCAGCTGCAGATATTAG	48787	Db	49784	CAGTGGAA-GGTTAGTGCAGAGATCTCAGGATTTATCAAT-GAGGCCATTGTCCCTCTATACC	49841
QY	1648	-AAAAAACTTCAAAAGTCTGCGTAGGCGCCGAGCAAACTTTAGAAACCTTATTGAACCTT	1706	QY	2722	CAGCTGTACCTAGGCCCTTATCTGTGMYTTCCTCCAAATAACGAGAGAAAGAGTGGTTTA	2781
Db	48788	AAAAAACTTCCAAAGTCGCTTAGGCGCGAGCAAAAGTTTAGAAACCTTACTGAACTT	48847	Db	49842	CAGCTGTACCTAACTCTTATATTCGCTTTCCTCCAAATACTAGAGAAAGCAAGTGGTTTA	49901
QY	1707	GCACACCTGCTTTTATATATAGATATCAGAGAGAGCGGCGGAACAGGACAAACGGGA	1766	QY	2782	CASCTCTGGACCTTMAAGATGCTTCTTCATCCCTCTGTATCATCTCCTGACTCTCAATCTCT	2841
Db	48848	GGCAACCTCGGTGTTTTTATATATAGATCAGAGAGAGCGGAGGAGCAAAATGGGA	48907	Db	49902	CAGTCTCTGGACCTTAAAGGATGCTTTTCTGCACCTCTATACATGCTGACTCTCAATCTCT	49961
QY	1767	T--TAAAAAAAGGCCACCGCTTTAGTCATGACCTCTCAGGCAAGTGGATTTGGAGGCTC	1824	QY	2842	TGTTTGCTTTGAAAGATACTTCAAAACCCARCATCTCAACTCAGCTGGACTTTTATACCCC	2901
Db	48908	TAAAAAAAAGGCGCACCTGCTTTAGTCATGCGCTCTCAGGCAAGCGGACTTTTGGAGGCTC	48967	Db	49962	TGTTTGCTTTGAAAGATCTTTCGAACCCAACTCTCAACTCAGCTGGACTTTTATACCCC	50021
QY	1825	TGGAAGAGGAAAGCTGGGCAAAATTGAATGCTTAATAGGGCTTGCTTCCAGTGGGCTCT	1884	QY	2902	AAGGGTTTCAGGGATAGYCCCATCTATTGTCAGGCAATGAGCCCAAGACTTTGAGYCAT	2961
Db	48968	TGGAAGAGGAAAGCTGGGCAAAATAGATGCTTAATAGGGCTTGCTTCCAGTGGGCTCT	49027	Db	50022	AAGGATTCAGGGATAGCCCCCATCTATTGTCAGGCAATGAGCCCAAGACTTTGAGCCAGT	50081
QY	1885	ACAAGGACATTTTAAAAAGATTGTCGAAGTAGAAGTAAGCCGCCCTTCTGTCATGCC	1944	QY	2962	YMTCATCTCTGGACACTCTTGTCTCTTCTGTAGTGGATGATTTACTTTTTCGCGCYRTT	3021
Db	49028	-CAAGGACATTTTAAAAAGATTATCCAAATAGAAATAAGCCCACTCCCTTGTCCATGCC	49086	Db	50082	TCTCATCTCTGGATATTCTTGTCTTGTGTATGCGGATGATTTACTTTTTCGCGCGCTT	50141
QY	1945	CTTATTTCAAGGAATCACTGAAGGCCCACTGCCCGCAGGAGCAAAAGTCTTTTGAGTC	2004	QY	3022	CAGAAACCTTGTGCGCATCAAGCCCAAGCRCTCTTMAATTTCTCGCYACTCTGGCT	3081
Db	49087	CTTATATCAAGGGAATCACTGTAAGGCCCACTTGCCCGCAGGAGCATGATGCTCTGAGTC	49146	Db	50142	CAGAAACCTTGTGCGCATCAAGCCCAAGGCTCTTAAATTTCTCGCCACTCTGGCT	50201
QY	2005	AGAAGCCACTAACACAGATGATCCAGCAGCAGACTGAGGGTCCCTGGGGCAAGCGCCATC	2064	QY	3082	ACAMGGTTTCCAAACSPARGCTCARTCTCTCTCAACAGCAG-----GTTAAA	3128
Db	49147	AGAAAGCCACTAACACAGATGATCCAGCAGCAGACTGAGGGTCCCTGGGGCAAGCGCCAGC	49206	Db	50202	ACAMGGTTTCCAAACSPARGCTCARTCTCTCTCAACAGCAGAGGGCTATTATCTCTATA	50261
QY	2065	CCATGCCATCACCCCTCACAGAGCCCTGGGTATGCTTGACCAATGAGGGCCAGGAAGT--	2122	QY	3129	TACTTAGGCTTAARATTTTCCAAAGGCACACGAGCCCTCAGTGAGGAAYRATCCAGCT	3188
Db	49207	CCATGCCATCACCCCTCACAGAGCCCTGGGTATGCTTGACCAATGAGGGCCAGGAAGTAA	49266	Db	50262	TACTTAGGCTTAARATTTTCCAAAGGCACACGAGCCCTCAGTGAGGAATGATCCAGCT	50321
QY	2123	-TGCTCTCGGACACTGGTGGCTCTTCTTACTCTTACTCTTCTCTGTCGCGCAACTGTC	2181	QY	3189	ATACTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRRTTCTCT-----	3239
Db	49267	CTGTCTCTCGGACACTGGTGGCTCTTCTTACTCTTACTCTTCTCTGTCGCGCAACTGTC	49326	Db	50322	ATACTGGCTTATCTCTTATCCCAAAACCTTAAACCAACTAAGAGGTTCTTGGCATATA	50381
QY	2182	CTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGGCAGTCACCTAGATATCTT	2241	QY	3240	-----GGCTAAAYAGVTTCTGCGAAWATGATTTCCCGAGTWTGCCAGTGGCAAATAGC	3290
Db	49327	CTCCATATCTGTCACTAT-----CCGAGCAGCCAGTCACCTAGATATCTT	49373	Db	50382	GGCATAAACAGGCATAACAGGTTTCTGCTGAATATGGAAT-CCCAAGTACGGCAAAATAGC	50440
QY	2242	CCGAGCCACTAAGTTATGAATCGGGAGCTTTATTTCTTTCACATGCTTTTCTAATTGTA	2301	QY	3291	CAGGYCATTAATATACASTAATTAAGGAACCTCAGAAAGCCATACCCATTTARTAGATG	3350
Db	49374	CCGAGCCACTAAGTTTGTG-ACTGGGGAACTTTTACTGTTTTTCAATGCTTTTCTAATTGTA	49432	Db	50441	CAGACCATTAATATACASTAATTAAGGAACCTCAGAAAGCCATACCCATTTTARTAGATG	50500
QY	2302	CTTGAAGCCCACTACCTTGTAGGAGAGACATCTAGCAAAAGCAGGGGCCATTTATA	2361	QY	3351	GAYAMCTCAAGYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCAAGYCCCA	3410
Db	49433	CTTGAAGCCCACTACCTTGTAGGAGAGACATCTAGCAAAAGCAGGGGCCATTTATA	49492	Db	50501	GACACTGAAGCAGGAGGAGCTTTCCAGGCCGTGAAGAAACACCTT--AACCCAGGCCCA	50558
QY	2362	CACCTGAACATAGGAGAGGAACACCCGTTGTTGTCGCTTGTAGGAAGGAATTA	2421	QY	3411	GTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATATYRTCACAGAAAAAACAGRAY	3470
Db	49493	CACCTGAACATAGGAGAGGAACACCCATTTGTTGT-CCCTTGTCTGAGGAAGGAATTA	49551	Db	50559	GTGTTAAGCTTCCAGCGGGCAAGACTTTTCTTCTATGTACACAG-AAAAAATAGGAAT	50617
QY	2422	TCTGAAAGTCTGGCAACAGAAAGGACATATGGAAGAGCAAAAGAAATGCCCTCTGTTTC	2481	QY	3471	AGCTCTGGAGTCTTTACAGACTCCAGGAGYAGCTTGCACACCTCTGAGGCTGACCTGAS	3530
Db	49552	TCTGAAAGTCTGGCAACAGAAAGGACATATGGAAGGACAAATACGGATGAG-CAAGAATGCCATCTGTTTC	49610	Db	50618	AGCTCTAGAGTCTTTACACAGGTCGAGGGACCAAGCTTGAACCCATCGCATCTAG	50677
QY	2482	AAGTTAACTAAAGGATTCCTTCTTCCCTACCAAGGAGGAGTACCCCTCAGACCA	2541	QY	3531	TAAAGAAATTCAGTGTAGTGGCAAGGGTTGRCYTCAATTGTTTAYGGGTAGTGGTGCAGT	3590
Db	49611	AAGTTAACTAAAGGATTCCTTCTTCCCTACCAAGGAGGAGTACCCCTTAGACCGG	49670	Db	50678	TAAAGAAATTCAGTGTAGTGGCAAGGGTTGCGCTCATTTGTTTACGGGTAGTGGGCGAGT	50737
QY	2542	AGGCCCAACAGGATTTCCAAAGATTGTTAAAGACTTAAAGCCCAAGGCTTAGTAAAC	2601	QY	3591	AGCAGTYKTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTTACTGTGTGGAC	3650
Db	49671	AGGCCCAACAGGATTTCCAAAGATTGTTAAAGACTTAAAGCCCAAGGCTTAGTAAAC	49730	Db	50738	AGCAGTYKTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTTACTGTGTGAAC	50797
QY	2602	CATGCAATACTCCCTGCAGTAATTCGTAGTGGATTGAGGAGGCAACAGAAACCCAGTGA	2661	QY	3651	ATCTCATGAGTGAAYRGCATCTACTGCTTAAAGGAGACTTGTGGCTGTGAGAACYCG	3710
Db	49731	CATGCAATACTCCCTGCAGTACTCC-----AATCTTACGAGTACAGAAACCCACAGA	49783	Db	50798	CTCTCATGATGTGAACCCGACATCTCAGTCTTAAAGAGAGCTTGTGGCTGTGAGAACCTG	50857
				QY	3711	TTTACTTAAATTCAGGCTCTATTACTTGAARGGCCAGTGTGTCGCRACCTGTGCACTTGTGC	3770



1408 ATTGCTGCTGACAGACATTTGCTAACTTGTGTGCTGTASAGGACTRAGGAACTASGAAGA 1467  
Db  
5696 ATTGTGCGCGACAGACATTTGCTAACTTGTGCTGTAGAAGGACTAAGGAAACTAGGAAG- 56754  
Qy  
1468 AARTCTAYGAATTTACTCAATGATGTCCACATAAACACAGGGGAAAGGAAAGAAATCCTAC 1527  
Db  
56755 AAGCCTATGAAATTTCAATGATGTCCACTATTAACACA-GGGAAGAAAGAAATCCTAC 56813  
Qy  
1528 TCCCTTTCTGGAGAGACTAAGGGAGGCATTTGAGGAAGCGTGCTCTCTGTCTACCTGACTC 1587  
Db  
56814 CACCTTTCTGGAGAGACTAAGGGAGGCATTTGACAAAGCATATCTCTGTCTACCTGACTC 56873  
Qy  
1588 TTCTGAAGCCAACTAACTTAAAGGTAAAGTTTATCACTCAGTCAGCTGACAGACATTAG 1647  
Db  
56874 TATTGAAGGCCAACTAACTTAAAGGAAAGTTTATCACTCAGTCAGCTGACAGATATTAG 56933  
Qy  
1648 -AAAAAACTTCAAAGTCTGCGGTAGGCCCGGAGCAAACTTTAGAAACCCCTATTGAACTT 1706  
Db  
56934 AAAAAAACTTCCAAAGTCGCTTAGGCCCGGAGCAAAAGTTAGAAACCTTACTGAACTT 56993  
Qy  
1707 GSCAACTCGGTTTTTATAATAGAGATCAGGAGGACGAGCGGAACAGGACAAACCGGA 1766  
Db  
56994 GGCAACCTCGGTTTTTATAATAGAGATCAGGAGGACGAGCGGAACAGGACAAATGGGA 57053  
Qy  
1767 T--TAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCAAGTGGACTTTGGAGGCTC 1824  
Db  
57054 TAAAAAAGGAGGCCACTGCTTTAGTCATGCGCCCTCAGGCAAGCGGACTTTGGAGGCTC 57113  
Qy  
1825 TCGAAAGGGAAGCTGGGCAAAATGGAATGCTTAATAGGCTTGCTCCAGTCGGGTCT 1884  
Db  
57114 TCGAAAGGGAAGCTGGGCAAAATAGAAATGCTTAATAGGCTTGCTCCAGTCGGGTCT 57173  
Qy  
1885 ACAAGGACACTTTAAAAAGATTTGTCACAGTAGAAGTAAGCGCCCTTCGTCATGCCC 1944  
Db  
57174 -CAAGGACACTTTAAAAAGATTTATCCAAATAGAAATAGCCACTCCCTTTGTCATGCC 57232  
Qy  
1945 CTTATTTCAAGGAATCACTGGAGGCCCACTGCGCCCGGAGGACAAAGTCTTTTGTAGTC 2004  
Db  
57233 CTTATATCAAGGGAATCACTGTAAAGGCCCACTGCCCCAGGAGACGTGAGTCTCTGAGTC 57292  
Qy  
2005 AGAAGCCACTAACCATGATATCCAGCAGCAGACTGAGGGTGTCTGGGGCAAGCGCATC 2064  
Db  
57293 AGAAGCCACTAACCATGATATCCAGCAGCAGACTGAGGGTGTCTGGGGCAAGCAGCAGC 57352  
Qy  
2065 CCATGCCATCACCTTACAGAGCCCTGGGTATGCTTGACCAATTTGAGGGCCAGGAAGGT-- 2122  
Db  
57353 CCATGCCATCACCTTACAGAGCCCTGGGTATGCTTGACCAATTTGAGGGCCAGGAAGGTAA 57412  
Qy  
2123 -TGCTCTGGACACATGCTGGGTCTTCTTTAGTCTTTTACTCTCTGTCTGTCGCGGACATGTC 2181  
Db  
57413 CTGTCTCTGGACACATGCTGGGTCTTCTCAGTCTTTACTCTCTGTCTGTCGCGGACATGTC 57472  
Qy  
2182 CTCAGATCTGTCACTATTCTGAGGGGTCCTTAAGACGGGAGTCACTAGATACTTTT 2241  
Db  
57473 CTCCTATCTGTCACTAT-----CCAGGACAGCCAGTCACTAGATAC--TTCT 57519  
Qy  
2242 CCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTTTCACTGCTTTTCTTAATATG 2301  
Db  
57520 CCCAGCCACTAAGTTGTG-ACTGGGGAACCTTTACTGTTTTCATGCTTTTCTTAATGTA 57578  
Qy  
2302 CTTGAAGGCCCACTACTTGTAGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATA 2361  
Db  
57579 CTTGAAGGCCCACTCTCTTGTAGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATA 57638  
Qy  
2362 CACTCAACATAGGAAGGAACACCCGCTTTGTTGTCNCCCTGCTTGGAGGAAGGAATTA 2421  
Db  
57639 CACTGAAATAGGAAGGAACACCCATTTGTTGT-CCCTGCTGGAGGAAGGAATTA 57697  
Qy  
2422 TCTTGAAGTCTGGGCAACAGGAACAATATGGACGAGCCAAAGAAATGCCCGTCTGTTTC 2481  
Db  
57698 TCTTGAAGTCTGGGCAACAGGAACAATATGGATGAG-CNAAGAAATGCCCATCTTGTTC 57756

2482 AAGTTAAACTAAAGGATTCACATCTCTTCCCTACCAAAAGCAGTACCCCTCAGACCCA 2541  
Db  
57757 AAGTTAAACTAAAGGATTCGCTCTCTTCCCTACCAAAAGCAGTACCCCTTAGACCGG 57816  
Qy  
2542 AGGCCCAACAAAGGATTTCCAAAGATTTGTTAAGGACTTTAAAAAGCCCAAGGCTTTAGTAAAC 2601  
Db  
57817 AGGCCCAACAAAGGACTCCAAAGATTTGTTAAGGACCTTAAAGCCCAAGGCTTAGTAAAG 57876  
Qy  
2602 CATGATTAACCTCCCTGCAAGTAATTCGCTAGTGGATTTAGGAGGACACAGAACCCAGTGA 2661  
Db  
57877 CATGCTAGTACCTCCCTGCAAGTATCC-----AAGTTTACGAGTACAGAAACCCCAACAGA 57929  
Qy  
2662 CAGTGGAGGTTAGTGAAGATCTCAGATTTATCAATGAGGCGCTTGTCTTTTATATACC 2721  
Db  
57930 CAGTGA-GGTTAGTGAAGATCTCAGATTTATCAAT-GAGGCCATTTGCTCTCTATACC 57987  
Qy  
2722 CAGCTGTACCTAGCCCTTATATCTGTGMYTTCCCAATAACAGAGGAAGCAGAGTGGTTTA 2781  
Db  
57988 CAGCTGTACCTAATCTCTTATATTCGCTTTCCCAATAACTAGAGGAAGCAAGTGGTTTA 58047  
Qy  
2782 CASTCTGGACCTTMAAGATGCTTCTGTGATCCCTGTATCATCTCAGCTCTCAATTC 2841  
Db  
58048 CAGTCTGGACCTTTAAGGATGCTTTTCTGCACCTCTATATCATGCTCAGCTCTCAATTC 58107  
Qy  
2842 TGTTCCTTTGAAGATATTTCAAACCCARCATCTCAACTCACCCTGGACTRTTTTACCC 2901  
Db  
58108 TGTTCCTTTGAAGATCTTTCGAACCCAACTCTCAACTCACCCTGGACTRTTTTACCC 58167  
Qy  
2902 AAGGTTTCAAGGATAGYCCCCATCTATTTGCGGAGCATTAAGCCCAAGCATTGAGYCAT 2961  
Db  
58168 AAGGTTTCAAGGATAGYCCCCATCTATTTGCGGAGCATTAAGCCCAAGCATTGAGCCAGT 58227  
Qy  
2962 YMTCACTGGACACTCTTGTCTGTCTGTAKGTGATGATTTACTTTTTCGCGCTTTC 3021  
Db  
58228 TCTCATACCTGGATATTTCTGTCTTTGGTATGCGGATGATTTACTTTTTCGCGCTTTC 58287  
Qy  
3022 CAGAAACCTTGGCCATCAAGCCCAAGGCTCTTMAATTTCTCGYACCTGTGGCT 3081  
Db  
58288 CAGAAACCTTGGCCATCAAGCCCAAGGCTCTTMAATTTCTCGYACCTGTGGCT 58347  
Qy  
3082 ACANGGTTTCAAGSAPARGCTCACTCTGTCTCAAGCAG-----GTTAAA 3128  
Db  
58348 ACAAGTTTCAAGSAPARGCTCACTCTGTCTCAAGCAGAGGCTTATTTATCTCTAAA 58407  
Qy  
3129 TACTTAGGCTTAARATTAATCAAGGACACAGGCGCTCAGTGAGGAAYATTCAGCT 3188  
Db  
58408 TACTTAGGCTTAARATTAATCAAGGACACAGGCGCTCAGTGAGGAAYATTCAGCT 58467  
Qy  
3189 ATACTGGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGGRTTCTCTT----- 3239  
Db  
58468 ATACTGGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGGRTTCTCTT----- 3290  
Qy  
3240 -----GGCTAAYAGVTTCTGCGGAATGATTTCCAGGTTGCGCAAAATAGC 3290  
Db  
58528 GGCATAACAGGATAACAGGTTTCTGCTGAATATGATTT-CCCAAGTACGCGCAAAATAGC 58586  
Qy  
3291 CAGGCTAATTAATTAATTAAGGAACTCAGAAAGCAATACCCATTARTATAGATG 3350  
Db  
58587 CAGACATTAATTAATTAAGGAACTCAGAAAGCAATACCCATTARTATAGATG 58646  
Qy  
3351 GAYAMCTCAAGYMAAGTGGCTTTCCAGGCGCTTAAAGAGGCTTTAAACCAAGYCCCA 3410  
Db  
58647 GACCTGAGCAGAGGAGCTTTCCAGGCGCTTAAAGAACCCCT--AACCCAGGCCCA 58704  
Qy  
3411 GTGTTAAGTTCGCAACRGGCAAGACTTTTSTTTATYRTTACAGAAACCAAGCAAGY 3470  
Db  
58705 GTGTTAAGTTCGCAACRGGCAAGACTTTTSTTTATYRTTACAGAAACCAAGCAAGY 58763  
Qy  
3471 AGCTCTRAGGCTTCTTACAGGCTCCRAGGAGYAGCTTTGCAACCTRTGCGYACCTGAS 3530  
Db  
58764 AGCTCTRAGGCTTCTTACAGGCTCCRAGGAGYAGCTTTGCAACCTRTGCGYACCTGAS 58823  
Qy  
3531 TAAAGAAATGTATGTAGTGGCAAGGTTGRCYCTATTGTTTAYGGGTAGTGGTGCAGT 3590

Db	58980	ATGAAGCAGATAGTACGAGGCCTGTGAAAGTTGTGCCAAAAAATAATCCCTTCGACTTCAGGC	59933
Qy	4670	CAAGC	4674
Db	59940	CATGC	59944
Search completed: August 1, 2006, 19:39:10 Job time : 5572 secB			

---

Db	58824	TNAGAAAATTGATGTAGTGCGAAGGGTTGGCCCTCATTTGTTTACGGGTAGTGGCGCAGT	58883
Qy	3591	AGCAGTYKTAGTATCTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGGAC	3650
Db	58884	AGCAGTCTTAGTATCTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGAAC	58943
Qy	3651	ATCTCATGKGTGAAYRGCACTACTCTGTAAGAGAGACTTGTGGCTGTGACAGAACVY	3710
Db	58944	CTCTCATGATGTGAACCGCATCTACTGCTAAAGAAGACTTGGTGGCTGTGACACAACVG	59003
Qy	3711	TTTTACTTAAATRTCAGGCTCTATTACTTTGAARGCCAGTGTGCRACCTGTGCACTTTGTGC	3770
Db	59004	TTTTGCTTAAATATCAGGCTCTATTACTTTGAGGGCCAGTGTGCTGACTGGCACCTTGTGC	59063
Qy	3771	AATCTTTAAACCAGYCNCAITTTCTTCCAGACAAATGAAGAAAAGATARAAYATACTGTCA	3830
Db	59064	AATCTTTAAACCAGGACATTTCTTCCAGACAAATGAAGAAAAGATAGAACAAGACTGTCA	59123
Qy	3831	ACAARTAAATTTCTCAAACCTATGCCACTCGAGGGGACCTTGTAGARGTTCCYTGTACTGA	3890
Db	59124	ACAAGTAAATTTGCTCAAACCTATGCCACTCGAGGGGACCTTCTAGTGTGTTCCCTTGTACTGA	59183
Qy	3891	TCCYGACCTTCAAACCTTGTATGATGGAAGTTCTTTGTAGAAAAGGACTTCGAAAAG	3950
Db	59184	TCCCAACC-TCAACTTTGTATGATGGAAGTTCTTTGTAGAAAAGGACTTCGAAAAG	59242
Qy	3951	YGGGGTATGAGTGGTCACTGATGAATGGAATYTTGAAAAGTAAATCCCCTCACTCCAGGAA	4010
Db	59243	CAGAGTGTGTAGTGGTCACTGATGAATGGAATACTTGAAGTAAATCCTCTGACTCCAGNA	59302
Qy	4011	CTAGTGTGTAGTRGCAGAACTAATAGCCVTCAATYTKGGGCACTAGAAATTTAGGAGAAGRAA	4070
Db	59303	CTAGTGTGTAGTRGCAGAACTAATAGCCCTCACTCAGGCACTAGAAATTTAGGAGAAGRAA	59362
Qy	4071	AAAGGVAAATATATATACAGACTCTRARTATGCTTACCTGCTTACCTGCTTCACTGCTCAGCAG	4130
Db	59363	AAAGGVAAATATATATACAGACTCTRARTATGCTTACCTGCTTACCTGCTTCACTGCTCAGCAG	59422
Qy	4131	CAATATGSAPAGAAAGGGAAATTCCTAACTTCYCAGRGAACACCTATCAMACATCAGGAAG	4190
Db	59423	CAATATGSAPAGAAAGGGAAATTCCTAACTTCYCAGRGAACACCTATCAMACATCAGGAAG	59482
Qy	4191	CMATTAGGAPATTTATTTAGGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGY	4250
Db	59483	CCATTAAGAAACTATTTATTTGGCTGCACAGAAATCTAAAGAGTGGCAGTCTTACACTGCT	59542
Qy	4251	GCGGTCTCATCANRAAGGAAGRAAGGAATASAGRGAAYTGCCNAGCAKATNTTGAAG	4310
Db	59543	G-----TAAGAAAGGACAGAGAAATAAAGGGAAACCCCGAGTGGATATTGAAG	59591
Qy	4311	CMAAAGAGCTGCAAGCGAGGACCTTCCATTTAGAAATGCTTATTAACCTTCCCTTTAGTAT	4370
Db	59592	CCGAAAGAGCCACNAGCGGACCTTCCATTTAGAAATGCTTATAGAAAGAACCGCTAGTAT	59651
Qy	4371	AGGGTAAATCCCTTTCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAA	4430
Db	59652	GGGGTAAATCCCTTTCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAA	59711
Qy	4431	CCTCAGGAGCAGTTTCTTCC-CCTCGGAGCGTTAGCCACTGAAGAGAGGAAATATCTT	4489
Db	59712	CCTCATGAGGACGTAGTTTCTTCCCTCCTCAGGATGGCTAGCCACCAAGAGGAAATATCTT	59771
Qy	4490	TTGCCCTGCAACTATCCAATGGAAATTTACTTTAAACCCCTTTCATCAACCTTTCCACTTAGGC	4549
Db	59772	TTGCCCTGCACTTACCAATGGAAATTTACTTTAAACCC-----TTCACTTAGGC	59819
Qy	4550	ATCGATAGCACCCATCARATGGCCAAATCATTTATTTACTGGACCGGCTTTTCAAACT	4609
Db	59820	ATTGATAGCACCCATCAGATGGCCAAATCATTTATTTACTGGACCGGCTTTTCAAACT	59879
Qy	4610	ATCAAGCARATKTGAGGCGCTGTGAATGTGCGCAPARAATAATCCCTCGCTTATCGC	4669

**This Page Blank (uspto)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:24:54 ; Search time 396 Seconds  
(without alignments)  
9633.402 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatcggtataaacc.....tattaatcttcgactgtr 7582

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1447667 seqs, 251571307 residues

Total number of hits satisfying chosen parameters: 2895334

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.New.\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	638.8	8.5	44063	US-10-539-228-718	Sequence 718, App
2	256.2	3.4	7323	US-10-539-228-118	Sequence 118, App
3	132.2	1.8	50000	US-10-528-659-2	Sequence 2, Appli
4	66.2	0.9	88493	US-10-539-228-703	Sequence 703, App
5	48.4	0.6	2579	US-11-218-305-4196	Sequence 4196, App
6	48.4	0.6	61791	US-10-539-228-645	Sequence 645, App
7	42.6	0.6	50	US-10-554-711-202	Sequence 202, App
8	42.2	0.6	6175	US-10-517-441-440	Sequence 440, App
9	42	0.6	3016	US-11-218-305-4195	Sequence 4195, App
10	41.6	0.6	50	US-10-554-711-192	Sequence 192, App
11	40.6	0.5	6175	US-10-517-441-714	Sequence 714, App
12	40.2	0.5	7975	US-11-256-173-12	Sequence 12, Appli
13	40.2	0.5	28605	US-10-471-571A-4495	Sequence 4495, App
14	39.2	0.5	1269	US-11-348-413-765	Sequence 765, App
15	37.6	0.5	2667	US-10-449-902-25901	Sequence 25901, A
16	37.6	0.5	155350	US-10-539-228-691	Sequence 691, App
17	37.2	0.5	913	US-11-174-307B-5089	Sequence 5089, App
18	36.8	0.5	1461	US-10-953-349-16839	Sequence 16839, A
19	36.8	0.5	1461	US-11-056-355B-59423	Sequence 59423, A
20	36.8	0.5	3943	US-10-449-902-16723	Sequence 16723, A
21	36.4	0.5	2040	US-11-348-413-6480	Sequence 6480, App
22	36.4	0.5	19634	US-10-517-441-688	Sequence 688, App
23	36	0.5	1153	US-11-218-305-7642	Sequence 7642, App
24	36	0.5	4944	US-11-218-305-1908	Sequence 1908, App
25	36	0.5	6001	US-10-517-441-772	Sequence 772, App

#### ALIGNMENTS

##### RESULT 1

US-10-539-228-718  
; Sequence 718, Application US/10539228  
; Publication No. US20060154250A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)  
; CURRENT APPLICATION NUMBER: US/10/539,228  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 718  
; LENGTH: 44063  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1) - (44063)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-718

Query Match 8.5%; Score 638.8; DB 6; Length 44063;  
Best Local Similarity 85.3%; Pred. No. 1e-185;  
Matches 732; Conservative 41; Mismatches 65; Indels 20; Gaps 6;  
QY 3690 CTTGGCTGTGTCAGAACGTTTAAATRTCAGGCTTATTCTTGAARGGCCAGT 3749  
DB 1 CTTGGCTGTGTCAGAACGTTTCTTAAATATCAGGCTTACTTGAAGGCCAGT 60  
QY 3750 GCTGCACTGTGCTTGTGCACTCTTAACCCAGCAGTCATTTCTCCAGACATGAAGA 3809  
DB 61 GTTGGCAATGCAATTTGTCAACTCT-----CAGCCATTTCTCTTCAACAT----- 110  
QY 3810 AAGGATARAAYATACTGTCAACAARTATTTCTCAACCTATGCCCTCGAGGGGACCT 3869  
DB 111 AAGACAGAATAACAATGTCAATAGGTGCTCAACCTATGTCTCGAGGGGACCT 170  
QY 3870 TTTAGAGTTCCCTTTGACTGATCCYGCACCTTCAACTTGTATCTGATGAAGTTCCTTTG 3929  
DB 171 TCTAGAGTTCCCTTTGACTGATCCCGACC-TCAACTTGTATCTGATGAAGTTCCTTTG 229  
QY 3930 TAGAAAAAGGACTTCGAAAGYGGGTATGCAAGTGGTCAAGTAAATGAATATTTGAAA 3989  
DB 230 TAGAAAAAGGACTTCGAAAGYGGGTATGCAAGTGGTCAAGTAAATGAATATTTGAAA 289

```
QY 3990 GTAATCCCTCACTCAGGAAGTAGTCTAGCTGCGAGAACTAATAGCCYTCATYKGG 4049
Db 290 GTAATCCCTCACTCAGGAAGTAGTCTAGCTGCGAGAACTAATAGCCCTCACTCGGA 349
QY 4050 CACTAGAATTAGGAGAGAAAGGGAATATATATATACAGACTCTTARTATGCTYACC 4109
Db 350 CACCAGAAATTAGGAGAGAAAGGGAATATATATATACAGACTCTTARTATGCTTACC 409
QY 4110 TAGTCNTCCATCCCATGMRGCAATATGSRAGAAGGGAATTCCTAACTTCYCGRGA 4169
Db 410 TAGTCCTCCATCCCATGCGCAATATGAGAGAAAGGGAATTCCTAACTTCGAGGGA 469
QY 4170 CACTATCMACATCAGGAAGCATTAGGARATTAATATGCGWGTACAGAAACCTARAG 4229
Db 470 CACTATCAACATCAGGAAGCATTAGGAGATTAATCTTGGCTCTACAGAAACCTAAG 529
QY 4230 AGTGGMAGCTTTACACTCYGGGGTCATCANAAAGGAAGRAAGGGAATASAGRA 4289
Db 530 AGTGGCAGCTTTACACTCCAGGGTCATCA-----GAAAGTAAAGGGAATAGAGGA 584
QY 4290 AYTGCCAAGCAKATATTAAGGMAAAGAGCTGCAAGGCGAGGACCTCCATTAGAAATGC 4349
Db 585 ACTGCCAAGCGGATATTGAAGCC-AAAGATCCGCAAGGCGAGGACCTCCATTAGAAATGC 643
QY 4350 TTATTAACTCCCTTAGTATAGGGTAATCCCTTCGCGGAAACCAAGCCCGAGTACTCG 4409
Db 644 TTATAGAAGGATCCCTTAGTATAGGGTAATCCCTTCGCGGAAACCAAGCCCGAGTACTCG 703
QY 4410 CAGGAGAAACAGAAATGGGAACTCAGCAGG--CAGTTTTCTCCCTCGGACGGTAG 4466
Db 704 CAGAGAAATAGAACGGGAACTCAGGAGACATAGTTTCTCCCTCAGGATGGCTAG 763
QY 4467 CCACGGAAGAGGAAATACATTTTGCCTGCAACTATCCAAATGGAATTAATTAAGCC 4526
Db 764 CCACGGAAGAGGAAATACATTTTGCCTGCGGCTAATCAATAGAAATTAATGTAAGGCC 823
QY 4527 TTCATCAACCTTTCACT 4544
Db 824 TTCACCAACCTTTCACT 841
```

## RESULT 2

```
US-10-539-228-118/c
; Sequence 118, Application US/10539228
; Publication No. US20060154250A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)
; CURRENT APPLICATION NUMBER: US/10/539,228
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US 10/322,281
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 73723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(73723)
; OTHER INFORMATION: n = A,T,C or G
US-10-539-228-118
```

```
Query Match 3.4%; Score 256.2; DB 6; Length 73723;
Best Local Similarity 45.4%; Pred. No. 6.1e-67;
Matches 1285; Conservative 83; Mismatches 1384; Indels 80; Gaps 17;
QY 2650 AAACCCAGTGGCAGTGGAGGGTTAGTGCAGAGATCTCAGGATTAATCAATGGAGCGGTG 2709
```

```
Db 58130 AAACCAGACAAGTCTTACAAGTTAGTTCAGGATCTCGCGCTTATGAA-CCAAATTAATTT 58072
QY 2710 TCCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTCCAAATACAGAGGAG 2769
Db 58071 TACTTATCCACCCGCTGGTGCAAAACCCATATACTCTTTTGTCTCAATACCTTCTCCA 58012
QY 2770 CAGAGTGGTTTACASTCCTCGACCTTMAAGGATCCTTCTGTGCATCCTGTACATCCTG 2829
Db 58011 CAATCACTGTTCGCTTTTGTATCTTAAAGATGCTTTTTTCACTATTCCTCGTCACT 57952
QY 2830 ACTCTCAATCTTGTGTGCTTTGAAGATACTT-----CAAACCCARCATCTCAAC 2880
Db 57951 TATCCCACTCTCTTTGCTTTTACCTGGACTGACCTCGACACCCATCAGTCCAGCAGT 57892
QY 2881 TCACCTGACATRTTTTACCCCAAGGGTTCAGGATAGVCCCATCTTATTTGGCCAGGAT 2940
Db 57891 TTACCTGGCTGTGTGCGCAAGGCTTTCAGGAGCAGCCCTCATTAATTTAGCCCAAGCTC 57832
QY 2941 TAGCCCAAGACTTGAGYCATYMTATACACTGGACACTTGTCTCTCGTAKGTGGATG 3000
Db 57831 TTTCTCATGATTTACTTTTCCACCCCTCCACTTCCACTTATCAATATATGATG 57772
QY 3001 ATTTACTTTTTCGCGCTTTTTCAGAAACCTTTGTGCATCAAGCCCAAGCRCTCTTMA 3060
Db 57771 ACCTTCTACTTTGTAGTCCCTCTTTGAGTCTTCTCAACAGACACCCCTCTGCTCTTC 57712
QY 3061 ATTTCTCGVACCTGTGGCTACAWGGTTTCCAAACARARGCTCARCTCTGCTCAGAC 3120
Db 57651 CCGTTACTGCTGTGGCATAATCTTTCATAAAACACACACTGCTCTCCCTGCCGATGTG 57592
QY 3181 TCCAGCTTATCTGCTTATCTCATCYCAAAACCCCTAA-AGCAACTAAGRRRTTCTTT 3239
Db 57591 TCCAAACTGATCTCTCAAAACCCCAACACTTCTCAAAACCAACACTCTTTCTTCTTA 57532
QY 3240 GGCRTAAYAGYTTCTGCCGAATATGATTTCCCGAGTGTGGCRAAATAGCAGGYCAT 3299
Db 57531 GGCATGGTTGGATACTTTCACTTTTGGATACCTG-GTTTGGCATCTCTAAACAACCCATT 57473
QY 3300 AWATACASTAATTAAGGAACCTCAGAAAGCCAAATACCCATTATTAAGATGGAYAMCTGA 3359
Db 57472 ATGTAATTTCAAAAAGGAAACCTAGCTGACCCCAT-----AGATCTAAATCTCT 57423
QY 3360 AGYMAAGTGGCTTTCCAGGCGCCCTTAAAGAGGCGCTTAAACCAAGYCCAGTGTAAGY 3419
Db 57422 TTCCCACTCTCTTTCCATTTCTTGAAGACAGATTTAGAGACTGCTCCACACTAGCTC 57363
QY 3420 TTGCCAACRGGGCAAGACTTTTSTTYYATYRTCAAGAAAAAAGCAGRAAY-AGCTCTRG 3478
Db 57362 TCCCTGACTCATCCCAACCCCTTTGGTTACACACAGCTGAAGTGCAGGGCTGTGCAGTTG 57303
QY 3479 GAGTCTTTACAGHTCCBAGGAYGAGCTTGCACACCTGTCRYACCTGASTAGGAAA 3538
Db 57302 GAATCTTTACACAGTACCGGACCGCGC-----CCTGTAGACCTTTTGTCAAACAAC 57249
QY 3539 YTGATGTAGTGGCAAGGGTTGRCYTCATTTTAYGGGTAGTGTGGCAGTAGCAGTYK 3598
Db 57248 TTGACCTTACTGTTTTTAGACTGGCCATGATGTCTCCATGCAGCGGCTGCCACTGACCTAA 57189
QY 3599 TAGTATCTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATG 3658
Db 57188 TACTTTTAGAGGCCCTTAAATCAACACTATGCTCAACTGCTGTACAGCTCTCTGTA 57129
QY 3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTTG---TGGCTGTACAGACAACGTTTAC 3715
Db 57128 ACTTCCAAAATCTATTTTCTTCTCCACTGTGATGATATATCTTCTGTCTCCCGAGCTCC 57069
QY 3716 TTAATATRCAGGCTCTATTAATTTGAARGCCAGTGTGCRATGTGCACTTGTGCAACTC 3775
Db 57068 TTCAGTTATCTCACTCTTTTGTGTGAGACTCCCACTTACCATTGTTCTTGTGCCCCAGACT 57009
```

Qy	3776	TTAACCCAGYCNCAATTTCTTCCAGACAATGAAGAAGAATARAAYATAACTGTGCAACAAR	3835
Db	57008	TCAAATCAGCGCTCCCACATTATTTCCGGATACCCACCTGACCCCCCATGCTGTATCTCTC	56949
Qy	3836	TAATTTTCTCAAACTGTGCCACTCGAGGGGACCCTVTAGARGTTTCYTTTGACTGATCGVG	3895
Db	56948	TGATCCACCTGCACATACACCCCATTTTCCCATGTCTTCTTCTTCTGTCTCTCACCTCTG	56889
Qy	3896	ACCTTTCAACTTGTATGATGAAGTTCCTTTGTAGAAAAAGAGACTTCGAAAAGYGGGG	3955
Db	56888	ATC-ACATTTGGTTTACTGATGCGCAGCTCCATCAGGCCTTAATCGCCACTCACCAAAG	56830
Qy	3956	TATGCACTGTGAGTGATATATGA---ATAVTTGAAAGTAATCCCCTCACTCCAGGA	4010
Db	56829	GCAGGCTATGCTATATATCTTCCACATCTATCATTTGAGGCTACCGGTCTGCCCTTACTCCA	56770
Qy	4011	CTAGTGTGYAGCTPRGCAGAACTAATAGCCYTCAYTKGGSCACTAGAAATTAGGAGAAGRAA	4070
Db	56769	CTACCTCTCAACAAGCGGAATCATTTGCCCTTAACTCGGSCCTCACTCTTGCAAAGGAT	56710
Qy	4071	AAGGGYAATATATATACAGACTCTPABTATGCTCYAACCTAGTCTNTCATGCCCATGMRG	4130
Db	56709	TACATGTCAATATTTATGCTGACTCPANAATAGCCTTCACATFCCTGCACCACTGCTG	56650
Qy	4131	CAATATGSARAGAAAGGGAATTCCTAACTCTCYAGRGAACACTATCAMACATCAGGAAG	4190
Db	56649	TTATATAGGCTGAAAGAGGTTTCTCACTACACAGAGTCTCTCAATCATTAATGTCTTTT	56590
Qy	4191	CCATTAGBARATTTATTAYTGGCWGTHACAGAAACCTARAGAGGTGGMAGTCTTTACACTGCY	4250
Db	56589	TAATAAAAACTCTCTCAAGGCGCTTTACTTCCAAAAGAAGCTGGAGTCAATTCACGTCA	56530
Qy	4251	GGGGTCAATCANAAAGGAAGRAAGGAAATASAGRGAAYTGCCACGAKANATTTGAAG	4310
Db	56529	AAGGCCATCAAAAGGATCAGATCCCATTTGCCCTCGGCAACGCTTATGCTGATAAGTAG	56470
Qy	4311	CMAAAGAGCTGCAAGGCGAGGACCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGPAT	4370
Db	56469	CTAAGNAGCAGTAG-----TGTTCCAACTTCTGTCTCCCTCAGCGCC	56428
Qy	4371	AGGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGGAAACAGAAATGGGAA	4430
Db	56427	AGTTTTTCTCTTCTCATCGGTCACTTCCACCTACTCCCCCACTGAACATC-----	56377
Qy	4431	CCTCAGGAGCACTTTTCTCCCTCGGAGGGTTAGCCACTGAAGAAGGGAATACTTTT	4490
Db	56376	CCACCTGTCAATCTCTTCCACATAGGCAAAATGGTTCTTAGACCAAGGAAGTATCTCC	56317
Qy	4491	TGCTCGCAACTATCCAAATGMAATTACTTAAACCTTCATCAAACTTTTTCACATTAGCCA	4550
Db	56316	TTCCAGCCTCAGAGGCCATTTATTGTGTGATCATTTTCATAACCTCTTCCATGTAAAGTT	56257
Qy	4551	TCGATAGCAACCATCARATGGCCAAATCAATTAATTTACTGGACCGAGCCCTTTTCAAAACTA	4610
Db	56256	ACAAGCGCTAGCCCGCCTCTTAGAATCTCTAATTTCTTTTCCATCATGGAATCTGTCC	56197
Qy	4611	TCAAGCARATAKTCAGGGCCTGTGAATGTTGCCARARAATAATCCCTCGCTTATCGCC	4670
Db	56196	TCAAGGAATCACTTCTCAGTGTGTTTCATCTGCTATTCTTACTACCCCTCAGGGATTTGTCA	56137
Qy	4671	AAGCTCTCT-----TCAGGARAAACAARAACAGGCCATTACCCTCGRAARACTGGC-AA	4723
Db	56136	GGCCCCCTCCCTTCCCTACATCAAGCTCAGGGACTTGGCCCCCTGCCAGGATTTGGCAAA	56077
Qy	4724	CTGAATTTATCCCAACAAGCCCCAAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATA	4783
Db	56076	TTGACTTTACTCAGATGCCCGAGTCAGGAACATAAAATCACTCTTGTCTGGGTAGACA	56017
Qy	4784	CTTTTCAGGGTGGCARAGGCCCTCCCTCTGTAGACAGAAAGGCCCAAGAGGTATATAA	4843
Db	56016	CTTTTCACTGGATGGGTAGAGGCCCTTTCCACAGAGGTCTGAGAAGGCCACCGTGTCTATTT	55957

QY	4844	AGGCACTAGTTTCATGAAATAATATCCCGAGTATCGGATCTTCCCGGAGGCTTACAGAGTGCACA	4903
Db	55956	CTTCCCTTCTGTGAGACATAAATCTCTCAGTTTGGCCCTTCCCACTCTATACAGTCCGATA	55897
QY	4904	ATAGCCCTCTCTTTCCAGGGCCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGAT	4963
Db	55896	ACGGATGAGCCCTTACTAGTTAAATTAACCAAGCAGTTTCTCAGGCTCTTGGTATTCAGT	55837
QY	4964	ATCATTACATCGCGCTGAAGGCCACAGTCTCTCAGGGAAGGTGCGAGAAAATGAATGAAA	5023
Db	55836	GGAACTCTTCAAAATCCCTTACCATCTCTCAATCTTTCAAGAAAGGTAGAACTAGACTAATGGTC	55777
QY	5024	YACTCAAAAGGACATCTTAAAAAAGCAAAACCCAGGAAACCCACCTTCACATGCGCTGTCGT	5083
Db	55776	TTTTAAAGACACACCTTCACAAGCTCAGCTTCCAAATTTAAAAGGACTGGACAGTACTTTT	55717
QY	5084	TGCTTATAGCTTTAAAAAGAAATCTGCAACTTTTCCCAAAAAG---CAGGACTTTAGCCCAT	5140
Db	55716	TACCTCTTGGCCCTTCTCAGAAATTAGAGCGCTGTCTCGAGATGCTACAGGGGTACAGTCCAT	55657
QY	5141	ACGAAATGCTGTATGGAAGGCCCTTCATA---ACCAATGACCTTTGTGCTTGACCCCAAGAC	5197
Db	55656	TTGAACTTTTATACGAGACGACCTTTCTTACTCAGCGCCCAACCTTGTGCCAGACACCGCC	55597
QY	5198	AGCCAACTTAGTTTGAGACATCACCTCTCTTAGGCCAAATATCAACAAGTCTCTTAAAAACATT	5257
Db	55596	CTCTAGGTGACTATCTCTCAGTCTCCAGCAGGCTA--GACAGGAATTCGCCAGGCTGC	55539
QY	5258	ACAAGGAACCTATCCCTGAGAAGAGGAAAAAGAACTATTTCCACCCMWGTGACATGGTATT	5317
Db	55538	TAATCTTCTTTGGCCCTACTCCAGATTCTCAACCATATGAACACACCCCTAGCTGGACAATC	55479
QY	5318	AGTCAAGTCCCTTCCTCTAAATTTCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCC	5377
Db	55478	AGTTCTTGTTAGAATACTGACCCCTCAAACCTCTACAACTCGATGGAGTGGACCCCTACTT	55419
QY	5378	AGTCAATTTTATVTAACCCCAACTGCGGTTAAAGTGCGCTGGAGTGGAGTCTTGGATACATCA	5437
Db	55418	AGTCACTATAGTATCCCAACTGCGGCTCCACCTGCGGATCCCTGAGGATCTCCCCCACTGGGTTACCG	55359
QY	5438	CACTTGAGTCAA	5449
Db	55358	TTCCAGAGGTAAA	55347
RESULT 3			
US-10-528-659-2/c			
; Sequence 2, Application US/10528659			
; Publication No. US200600995941			
; GENERAL INFORMATION:			
; APPLICANT: NAGATA INDUSTRIAL SCIENCE RESEARCH INSTITUTE			
; APPLICANT: GIFU INTERNATIONAL INSTITUTE OF BIOTECHNOLOGY			
; APPLICANT: YAMADA, Yoshiji			
; APPLICANT: YOKOTA, Mitsuhiro			
; TITLE OF INVENTION: Method for diagnosing a risk of hypertension			
; FILE REFERENCE: C0200701			
; CURRENT APPLICATION NUMBER: US/10/528,659			
; CURRENT FILING DATE: 2005-03-22			
; PRIOR APPLICATION NUMBER: JP P2002-280034			
; PRIOR FILING DATE: 2002-09-25			
; NUMBER OF SEQ ID NOS: 36			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 2			
; LENGTH: 50000			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-528-659-2			

Query Match 1.8%; Score 132.2; DB 6; Length 50000;  
Best Local Similarity 44.2%; Pred. No. 1.2e-28;  
Matches 730; Conservative 71; Mismatches 797; Indels 54; Gaps 11;  
Qy 2706 GTTGTCTCTTTTATCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAATATCCAGAG 2765



Db 66649 ACTGAGGCTTTCCCAACCAACATGACACCATTCATAGTGACAAAGAGCTCTTAGAG 66590  
QY 4858 GAATATATCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTC 4917  
Db 66589 GACATAGTGCACAGGTATGGGCTTCTCTCA--GATGATAGTCAGACAATGGACCAGCTCTT 66532  
QY 4918 CAGGCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGC 4977  
Db 66531 GTGCCCCAGGTAGTCAGAGCTAGTCAATATCTCGAGATGATTAATAACT-----TGT 66476  
QY 4978 GCTTGAAGGCCACAGTCTCTCAGGAAAGTTCGAGAAATGAATGAAYACTCAAGGACAT 5037  
Db 66475 ATATATAGACCCAGAGTTCAGGACAGGTAGAAAAATAAACAAATCTTAAGACCC-- 66418  
QY 5038 CTAATAAAGCAACCCAGGAACCCACCTCACATGGCTGCTCTGTTGCTCTATAGCCTTA 5097  
Db 66417 TTAACCTAACTGACCTTAAGAGACTAGTGACTGAATAACTCTCTTCCCTTTACTCTC 66358  
QY 5098 AAAAGAAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGA 5157  
Db 66357 TATAGGCGTGAATCTCCCTTACCAGATGAACTACCCCTTTTAAATAATATGTTGCGG 66298  
QY 5158 AGGCCCTTC 5166  
Db 66297 ATACCACTC 66289

## RESULT 5

US-11-218-305-4196  
; Sequence 4196, Application US/11218305  
; Publication No. US20060141495A1  
; GENERAL INFORMATION:  
; APPLICANT: MONSANTO TECHNOLOGY, LLC  
; APPLICANT: McInaird, Paul L.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Wu, Kuneheng  
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping  
; TITLE OF INVENTION: Corn.  
; FILE REFERENCE: 38-21 (53660) B  
; CURRENT APPLICATION NUMBER: US/11/218,305  
; PRIOR FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US 60/606,880  
; PRIOR FILING DATE: 2004-09-01  
; NUMBER OF SEQ ID NOS: 25043  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4196  
; LENGTH: 2579  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-11-218-305-4196

Query Match 0.6%; Score 48.4; DB 7; Length 2579;  
Best Local Similarity 45.5%; Pred. No. 0.0012;  
Matches 160; Conservative 3; Mismatches 189; Indels 0; Gaps 0;  
QY 4746 ACTCAGGGATTTCAGTATCTAGTCTGGGTATATCTTCCCGGTTGGCAGAGGC 4805  
Db 1378 ATCAAGGGGCATCGGTGTGTGTAGTTCCTGACTACTTCCACCAATGGACTGAGC 1437  
QY 4806 CTTCCCTCTAGCAGAGAAAGCCAGAGGTAAATAGGCACTAGTTCATGAAATAT 4865  
Db 1438 CGTGTCTTAAGAACATGACACACAGGAGGTAAATGATTTTATACTGAGCATATTAT 1497  
QY 4866 TCCAGATTCCGACTTCCCGAGGCTTACAGATGACATAGCCCTTCCAGGCCAC 4925  
Db 1498 TCATAGATTCCGATTTCCCGAGACCTTAATCAGATCAAGGACTCTTTTATGTCAA 1557  
QY 4926 AGTAACCCAGGGAGTATCCAGCGGTTAGGTATACGATATCACTTACATCGCCCTGAAG 4985  
Db 1558 GGAGGTACGTGAATTTGCTGAATTTATACAGAAATTAAGCTCTTAATTCATCTCATTA 1617  
QY 4986 GCCACAGTCTCAGGGAAGTCGAGAAATGAATGAAYACTCAAGGACATCTAAAAA 5045

Db 1618 TGCTCAGGCCAATGGACAGCCGAGTCTAGTAATAGGACATGATTAAATTTGATAAAAA 1677  
QY 5046 GCAAAACCCAGGAACCAACCTCACATGGCCTGTCTGTTGCCCTATAGCCTTA 5097  
Db 1678 GAAGATATCTGATAATCTTAACATTGGCATAAGATTTTCTCTGAAGCTTTA 1729

## RESULT 6

US-10-539-228-645/c  
; Sequence 645, Application US/10539228  
; Publication No. US20060154250A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)  
; CURRENT APPLICATION NUMBER: US/10/539,228  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 645  
; LENGTH: 61791  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(61791)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-645

Query Match 0.6%; Score 48.4; DB 6; Length 61791;  
Best Local Similarity 51.2%; Pred. No. 0.014;  
Matches 125; Conservative 5; Mismatches 111; Indels 3; Gaps 1;

QY 6614 AGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGTCCCGACTCCC 6673  
Db 58689 AGAATTACAGAACCATATATACATTGCACTAGATAGCAATATAGAGAAATAGACAAGTCTA 58630  
QY 6674 TGCTCACCCTGCAAGATCAACTTACTCCCTAGCAGCAGTAGTCTTCRAAATCGAAGAG 6733  
Db 58629 TTAATGCTCTACAAGATTCCCTTTCTCTGCTGAAACAGTAAACCAATTGAAGAG 58570  
QY 6734 CTTTAGACTYGTCTAACCGCTGARAGAGGGGACCTGTTTATTTTATTTAGGGGAAGATGCT 6793  
Db 58569 GATTGGAGTTGTT---TCTTCAGCAGGGTGGTTTATGTTTTCAGCTCTTGGAGAAGATGCT 58513  
QY 6794 GTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCGAA 6853  
Db 58512 GTGCTTATGCTGCTCACTCCAAAGTAGTCAAGAGATATGGCTCAGCTTAGAAAAAGAT 58453  
QY 6854 TACA 6857  
Db 58452 TAAA 58449

## RESULT 7

US-10-554-711-202  
; Sequence 202, Application US/10554711  
; Publication No. US20060115806A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Georgia Research Foundation  
; APPLICANT: McDonald, John F.  
; TITLE OF INVENTION: Global Analysis of Transposable Elements  
; FILE REFERENCE: as Molecular Markers of Cancer  
; FILE REFERENCE: 21099.0075P1  
; CURRENT APPLICATION NUMBER: US/10/554,711  
; CURRENT FILING DATE: 2005-10-28  
; PRIOR APPLICATION NUMBER: 60/466,798  
; PRIOR FILING DATE: 2003-04-29  
; NUMBER OF SEQ ID NOS: 778

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: Synthetic Construct
US-10-554-711-202

Query Match          0.6%; Score 42.6; DB 6; Length 50;
Best Local Similarity 91.8%; Pred. No. 0.0035;
Matches 45; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7127 AAGGCACCCCTCCTGAGGAAATCTCAGCTGCAACCTCTACTAGGCC 7175
      |||||
Db 2 AAGGCACCCCTCCGAGGAAATCTCAACTGCACGACCCCTACTAGGCC 50

RESULT 8
US-10-517-441-440/c
; Sequence 440, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 440
; LENGTH: 6175
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-440

Query Match          0.6%; Score 42.2; DB 6; Length 6175;
Best Local Similarity 44.7%; Pred. No. 0.19;
Matches 98; Conservative 14; Mismatches 107; Indels 0; Gaps 0;

QY 3198 TATCCTCATCYCAAAACCCCTAAGCACTAAGRGRTTCCTTGGCRTAAYAGGYTTCTGC 3257
      |||||
Db 5675 TTCTTAACCACTACCAAAACGCAAAACAACTTCTATATAATTAATTACTTCTTC 5616

QY 3258 CGAATATGATTTCCCGAGTWTGGCRAATAGCCAGGYCATTAATACASTAATTAAGGA 3317
      |||||
Db 5615 TTATAAAAAACCCATTTAATAATACCTTATAAATCAAAACCTTAATAAAACAATACTTAAATA 5556

QY 3318 AACTCAGAAGCAATACCATTTTARTAGATCGAVAMCTGAAGYVRAAGTGGCTTCCA 3377
      |||||
```

```
Db 5555 TAATCAACACACAAAAAACTATTCAAACTATATAATACTAAATCCAACTCGTATCACT 5496

QY 3378 GGCCCTTAAGAGCGCTTAAACCCAAAGYCCAGTGTTA 3416
      |||||
Db 5495 TACCACACACACCAATAAATCTAAAAACAATAATTA 5457

RESULT 9
US-11-218-305-4195
; Sequence 4195, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: Mc Laird, Paul L.
; APPLICANT: Tao, Kunsheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4195
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-4195

Query Match          0.6%; Score 42; DB 7; Length 3016;
Best Local Similarity 44.3%; Pred. No. 0.13;
Matches 156; Conservative 3; Mismatches 193; Indels 0; Gaps 0;

QY 4746 ACCTCAGGATTTTCAGTATCTACTGCTGGTARATATCTTTTCAGGGTTGGGCARAGC 4805
      |||||
Db 1998 ATCAAGGGCATCGTTGCTGTTAGTTGCCACTGACTACTTACCAATGAGCTGAAGC 2057

QY 4806 CTTCCCTGTAGGACAGAAAGCCCAAGAGGTATAAAGGCATAGTTTCATGAATAAT 4865
      |||||
Db 2058 CGTTGCTCTGAAGAACATGACACACAGGAGGTAAATGAGTTCAATACTGACATATTAT 2117

QY 4866 TCCAGATTCGAGCTTCCCGAGGCTTACAGAGTACAGATACCTGCTTCCAGGCCAC 4925
      |||||
Db 2118 TCATAGATTCGGCATTTCCCGAGACCTTCACACTCAAGGACTTCTTTTATGTCAA 2177

QY 4926 AGTAACCCAGGAGTATCCCGGCTTAGGTATACGATATCACTTACACTGGCGCTGAAG 4985
      |||||
Db 2178 GGAGTACGTGAATTTGCTGAATTAATACAGAAATTAAGCTGCTTAATTCATCTCCATATTA 2237

QY 4986 GCCACAGTCTCAGGAAAGGTCGAGAAATGAATGAAYACTCAAAAGGACATCTAAAAA 5045
      |||||
Db 2238 TGCTCAGGCAATGGACAGGCGAGCTAGTAATAGGACATTGATTAAATTTGATAAAAA 2297

QY 5046 GCAAACCCAGGAAACCCACCTCACATGGCTGCTGTTGCTCCTATAGCCTTA 5097
      |||||
Db 2298 GAAGATATCTGATAATCTCTAAACATTGGCATAAGATTTTGTCCGAAGCTTTA 2349

RESULT 10
US-10-554-711-192
; Sequence 192, Application US/10554711
; Publication No. US20060115806A1
; GENERAL INFORMATION:
; APPLICANT: University of Georgia Research Foundation
; APPLICANT: McDonald, John F.
; TITLE OF INVENTION: Global Analysis of Transposable Elements
; TITLE OF INVENTION: as Molecular Markers of Cancer
; FILE REFERENCE: 21099.0075P1
; CURRENT APPLICATION NUMBER: US/10/554,711
; CURRENT FILING DATE: 2005-10-28
; PRIOR APPLICATION NUMBER: 60/466,798
```

```
; PRIOR FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 778
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: Synthetic Construct
US-10-554-711-192

Query Match      0.6%; Score 41.6; DB 6; Length 50;
Best Local Similarity 88.0%; Pred. No. 0.0072;
Matches 44; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 456 CAACGAACTCCAGTCAGAGAATACGARGCTTGCACCATCTTGGGAAGCG 505
    |||||
Db 1 CCAAGAACCCAGGTCAGAGAACACGAGGCTTGCACCATCTTGGGAAGTG 50

RESULT 11
US-10-517-441-714/c
; Sequence 714, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOKKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFLE, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 714
; LENGTH: 6175
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-714

Query Match      0.5%; Score 40.6; DB 6; Length 6175;
Best Local Similarity 44.3%; Pred. No. 0.6;
Matches 97; Conservative 14; Mismatches 108; Indels 0; Gaps 0;

QY 3198 TATCCTCATCYGAAACCCCTAAGCAACTAAGRGRTTCTTGGCTTAAYAGGYTTCTGC 3257
    |||||
Db 5675 TTCTTAAACCACTACCAAAAACACAAAAAACTTTCTATAATAATAATTACTTCTTC 5616

QY 3258 CGAAATGATTTCCCGAGGTGTCGCAATAGCCAGGTCATTATATACATACTAATTAAGCA 3317
    |||||
Db 5615 TTATAAAAAACCCATTAAATAACCTATAAATCAACCTAATAAATAAATACTTAATA 5556

; PRIOR FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 778
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: Synthetic Construct
US-11-256-173-12

Query Match      0.5%; Score 40.2; DB 7; Length 7975;
Best Local Similarity 46.0%; Pred. No. 0.98;
Matches 93; Conservative 10; Mismatches 99; Indels 0; Gaps 0;

QY 3970 TGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTCTVAGCTRCAGA 4029
    |||||
Db 7089 TAATATGCGAAACATGCAATTAATGTTAGCGAAACTTAAACAATGCCAAACACGACG 7148

QY 4030 ACTAATAGCCYTCAYTKGGGCACATAGAATTAGGAGAGRAAAAGGYYAAATATATATAC 4089
    |||||
Db 7149 GATTACAGCAATCAATGGCGCATCTGATTTAAATCAAAACAAAGAGATGCTTAAAGC 7208

QY 4090 AGACTCTTRATATGCTTACCTAGTCTGATGTCATGCCCATGMRGCAATATGSAAGAAAGGA 4149
    |||||
Db 7209 ACAAGCTATGCTGCTCAACGGGTATCTAATGCAAGATGTACAGCACAATGCGACTGA 7268

QY 4150 ATTCTTAACCTTCVAGRGACA 4171
    |||||
Db 7269 ACTGAACACGCAATGGGCACA 7290

RESULT 13
US-10-471-571A-4495
; Sequence 4495, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4495
; LENGTH: 28605
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-4495

Query Match      0.5%; Score 40.2; DB 6; Length 28605;
Best Local Similarity 46.0%; Pred. No. 2.6;
Matches 93; Conservative 10; Mismatches 99; Indels 0; Gaps 0;
```



[illegible]

```

RESULT 14
US-11-348-413-765
; Sequence 765, Application US/11348413
; Publication No. US20060160121A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348.413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 765
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1269)
; OTHER INFORMATION: WAN01P6H9
; OTHER INFORMATION: :SPY0907:pyrC:putative dihydrootase:SF370:NC_002737.1
US-11-348-413-765

```

	Query Match	0.5%	Score 39.2;	DB 7;	Length 1269;
	Best Local Similarity	52.3%;	Pred. No. 0.48;		
	Matches 69;	Conservative 7;	Mismatches 56;	Indels 0;	Gaps 0;
Qy	3513	ACCYRGGCRVACCTGASTAAGGAAATGATGTAGTGGCAAAGGTTGRCYCTCATTCGTTT	3572		
Db	21	ACCGGTTATGGACCCGAAATCGCARAGAGATCAAGTGGCAGATGCTCTTATTGATGGTAA	80		
Qy	3573	AVGGGTAGTGGTGCCAGTACGATYKTAGTATCTGAAGCAGTTTAAAAATAATACAGGGRAG	3632		
Db	81	ACAGATTGTAAAGATAGCTTCACGCTATTGAATGCCAAGAAGCTCAAGTGATTGATGCTAG	140		
Qy	3633	AGATCTTACTGT	3644		
Db	141	TGGCTCTATTGT	152		

RESULT 15  
US-10-449-902-25901  
; Sequence 25901, Application US/10449902  
; Publication No. US20060123505A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agrobiological Sciences.  
; APPLICANT: Bio-oriented Technology Research Advancement Institution.  
; APPLICANT: National Institute of Agrobiological Sciences.

```

; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25901
; LENGTH: 2867
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK101343
; DATABASE ENTRY DATE: 2002-08-28
; US-10-449-902-25901

```

	Query Match	0.5%	Score 37.6;	DB 6;	Length 2667;
	Best Local Similarity	48.1%	Pred. No. 2.7;		
	Matches 100;	Conservative 1;	Mismatches 107;	Indels 0;	Gaps 0;
Qy	130	TCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGCTGTTTGCCA	189		
Db	12	TCTTTTTTCGCCAATCTCGATCGCTTGTCTCCCTCCCGTGGGCTCCGAT	71		
Qy	190	CCACCGCANACCTGCCGCTGACTCCCATCCTCTGGATCTCGAGGTTGCCGTGTGCT	249		
Db	72	CCGCGGCCAATCCGCCCGGAATCCGCGGAATCCGGCGGGGCAACGGCGCGCGCGCG	131		
Qy	250	CCTGATCCAGGARGCGCCCATTTGCCGCTCCCAATTGGGCTTAAAGGCTTGCCATTGTNCC	309		
Db	132	ATTGTGTGGGAGGCGCGCGCGGCGCTCGGATCGAGTCCGGGTGACGGATTGGC	191		
Qy	310	TGCACGCTAAGTGCCCTGGGTTTTGTCT	337		
Db	192	CGGTTGGAAACGGGAGCTTCGACTCGATT	219		

Search completed: August 1, 2006, 21:03:46  
Job time : 403 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:10:51 ; Search time 30582 Seconds  
(without alignments)  
13863.723 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7489.6  
Sequence: 1 caacaatcggtataaacc.....tattaaattctgcactgtg 7582

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est3: \*  
3: gb\_est4: \*  
4: gb\_est5: \*  
5: gb\_est6: \*  
6: gb\_hic: \*  
7: gb\_est2: \*  
8: gb\_est7: \*  
9: gb\_est8: \*  
10: gb\_est9: \*  
11: gb\_gss1: \*  
12: gb\_gss2: \*  
13: gb\_gss3: \*  
14: gb\_gss4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180.4	29.1	2748	6	CR605851
2	2176.4	29.1	2749	6	CR617248
3	2174.4	29.0	2500	6	CR622175
4	2169.4	29.0	2748	6	CR625046
5	2147.4	28.7	2716	6	CR613169
6	969	12.9	1071	4	CR636506
7	921.2	12.3	998	4	CR637769
8	909.6	12.1	1058	4	CR637830
9	908.6	12.1	935	4	CR639174
10	885.2	11.8	955	4	CR634795
11	885	11.8	1019	4	CR643963
12	866.6	11.6	1014	4	CR637830
13	848.8	11.3	6400	6	CR749521
14	834.8	11.1	924	4	CR634731
15	820.8	11.0	931	4	CR638965
16	820	10.9	903	4	CR640875
17	812	10.8	872	4	CR632664
18	810.4	10.8	6427	6	CR857090
19	804	10.7	883	4	CR640873

c	20	803.8	10.7	966	4	BX380176	BX380176
c	21	789.2	10.5	877	4	BX347619	BX347619
c	22	776.2	10.4	921	4	BX408734	BX408734
c	23	771	10.3	836	4	BX388766	BX388766
c	24	768.6	10.3	995	4	BX389657	BX389657
c	25	767.6	10.2	808	9	DA834147	DA834147
c	26	765.4	10.2	844	4	BX347111	BX347111
c	27	764.8	10.2	828	4	BX367907	BX367907
c	28	753.8	10.1	846	4	BX368078	BX368078
c	29	751.8	10.0	820	9	DA833993	DA833993
c	30	751.6	10.0	900	4	BX459153	BX459153
c	31	734.2	9.8	836	13	CZ457645	CZ457645
c	32	732.6	9.8	770	9	DR004257	DR004257
c	33	730.6	9.8	769	9	DA833923	DA833923
c	34	724.6	9.7	758	4	BX357208	BX357208
c	35	723.4	9.7	879	4	BX325902	BX325902
c	36	719.4	9.6	924	4	BX409328	BX409328
c	37	718.6	9.6	761	9	DA832031	DA832031
c	38	715	9.5	782	9	DA833912	DA833912
c	39	709.4	9.5	771	2	BI087886	BI087886
c	40	704.2	9.4	792	9	DA830020	DA830020
c	41	700.4	9.4	840	9	DA834086	DA834086
c	42	697.2	9.3	721	1	AU138405	AU138405
c	43	686.2	9.2	819	9	DA830049	DA830049
c	44	679.4	9.1	799	13	CZ450650	CZ450650
c	45	678.2	9.1	832	14	DU836611	DU836611

## ALIGNMENTS

RESULT 1  
CR605851  
LOCUS  
DEFINITION  
full-length cDNA clone CS0DE012YJ24 of Placenta of Homo sapiens (human);  
ACCESSION  
CR605851  
VERSION  
CR605851.1 GI:50486658  
KEYWORDS  
HTC; CNSLT cDNA.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 2748)  
AUTHORS  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished  
REMARK  
Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
REFERENCE  
2 (bases 1 to 2748)  
AUTHORS  
Genoscope.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
FEATURES  
Location/Qualifiers  
source  
1..2748  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE012YJ24"  
/tissue\_type="Placenta"  
/plasmid="pCMVSPORT\_6"  
ORIGIN  
Query Match 29.1%; Score 2180.4; DB 6; Length 2748;

Best Local Similarity 98.4%; Pred. No. 0;			
Matches 2171; Conservative 34; Mismatches 0; Indels 1; Gaps 1;			
QY	5352	GATACATCTCGGAAGAGCCCTACCCAGTCATTTTATYATCCCAACCTGCGGTAAAGTG	5411
Db	544	GATACATCTCGGAAGAGCCCTACCCAGTCATTTTATCTACCCCAACTGCGGTAAAGTG	603
QY	5412	GCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAATCTGATGCTGCAAGGAA	5471
Db	604	GCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAATCTGATGCTGCAAGGAA	663
QY	5472	CCTGAAATCCAGGAGACAACTCTAGCTATCTCTGTAACCTCTAGAGGATTTGCGCTG	5531
Db	664	CCTGAAATCCAGGAGACAACTCTAGCTATCTCTGTAACCTCTAGAGGATTTGCGCTG	723
QY	5532	CTCTTCAAACAACCCAGGAGAAAGTAACATAAATCAATAATCCCCATGSGSCCTCCC	5591
Db	724	CTCTTCAAACAACCCAGGAGAAAGTAACATAAATCAATAATCCCCATGSGSCCTCCC	782
QY	5592	TTATCATATTTTCTCTASTGTTSTTTTACCTSTTTTCACTCTCACTGCAACCCCTCC	5651
Db	783	TTATCATATTTTCTCTASTGTTSTTTTACCTSTTTTCACTCTCACTGCAACCCCTCC	842
QY	5652	ATGCGCGCTGTATGACCACTAGCTCCCTVACCMAGAGTTTCTATGGAGAAATGCAGCGTCC	5711
Db	843	ATGCGCGCTGTATGACCACTAGCTCCCTTACCAAGAGTTTCTATGGAGAAATGCAGCGTCC	902
QY	5712	CGGAAATTTGATGCCCCCATGTATAGGAGTCTTTSTAAGGAAACCCCACTTCACTGC	5771
Db	903	CGGAAATTTGATGCCCCCATGTATAGGAGTCTTTSTAAGGAAACCCCACTTCACTGC	962
QY	5772	CCACACCATATGCCCGCAACTGCTATCACTCTGCCACTTTTCATGTCATGCAATAC	5831
Db	963	CCACACCATATGCCCGCAACTGCTATCACTCTGCCACTTTTCATGTCATGCAATAC	1022
QY	5832	TCATTTATGGACAGGAAATGATTAATCTTAGTTGCTGCGAGACTTGGAGTCACTGT	5891
Db	1023	TCATTTATGGACAGGAAATGATTAATCTTAGTTGCTGCGAGACTTGGAGTCACTGT	1082
QY	5892	CTGTTGGACTTACTTCAACCACTGGTATGCTGATGGGGTGGAGTTCAAGATCAGGC	5951
Db	1083	CTGTTGGACTTACTTCAACCACTGGTATGCTGATGGGGTGGAGTTCAAGATCAGGC	1142
QY	5952	AGAGAAAAACATGTAAAGAGTAATCTCCAACTCACCCSGGTACATGGCACTCTAG	6011
Db	1143	AGAGAAAAACATGTAAAGAGTAATCTCCAACTCACCCSGGTACATGGCACTCTAG	1202
QY	6012	CCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCATCTCGCT	6071
Db	1203	CCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCATCTCGCT	1262
QY	6072	GGTAAGCCTATTAAATACACCTCTCACTGGGTCCATAGGTCTCGGCCAAAACCCCTAC	6131
Db	1263	GGTAAGCCTATTAAATACACCTCTCACTGGGTCCATAGGTCTCGGCCAAAACCCCTAC	1322
QY	6132	TAACTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACTGA	6191
Db	1323	TAACTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACTGA	1382
QY	6192	ACAAATGGAACAATTCAGACAGAAATAAACCACTTCCTGTTTATAGGACCTCTTGT	6251
Db	1383	ACAAATGGAACAATTCAGACAGAAATAAACCACTTCCTGTTTATAGGACCTCTTGT	1442
QY	6252	TTCCAAATSTGGAATAACCCATACCTTCAAACTCTCACTGTGTAAATTTAGCAATCTAC	6311
Db	1443	TTCCAAATSTGGAATAACCCATACCTTCAAACTCTCACTGTGTAAATTTAGCAATCTAC	1502
QY	6312	ATACAAACAACCTCCCAATGATCAGGTGGGTAACTCTCCCAACAAATAGTCTGCT	6371
Db	1503	ATACAAACAACCTCCCAATGATCAGGTGGGTAACTCTCCCAACAAATAGTCTGCT	1562
QY	6372	ACCCTCAGGAATATTTTGTCTGTGTTACCTCAGCCTATCGTTGTTGAATGGCTCTTC	6431
Db			
Db	1563	ACCCTCAGGAATATTTTGTCTGTGTTACCTCAGCCTATCGTTGTTGAATGGCTCTTC	1622
QY	6432	AGAATCTATGCTTCTCTCTCATTTTAGTCCCGCCCATGCACTTACACTGACACAGA	6491
Db	1623	AGAATCTATGCTTCTCTCTCATTTTAGTCCCGCCCATGCACTTACACTGACACAGA	1682
QY	6492	TTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTTGT	6551
Db	1683	TTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTTGT	1742
QY	6552	TATAGGAGCAGGAGTGTAGTGCACTAGTATCTGGCAATTCGCGGTATCACAACCTCTAC	6611
Db	1743	TATAGGAGCAGGAGTGTAGTGCACTAGTATCTGGCAATTCGCGGTATCACAACCTCTAC	1802
QY	6612	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGGACTC	6671
Db	1803	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGGACTC	1862
QY	6672	CCTGCTCACTTGCAGAGTCAACTTAATCTCCTAGCAGCAGTAGTCTTCAAATTCGAAG	6731
Db	1863	CCTGCTCACTTGCAGAGTCAACTTAATCTCCTAGCAGCAGTAGTCTTCAAATTCGAAG	1922
QY	6732	AGCTTTAGACTTGTAAACCGCTGARAGAGGGGAACTGTTTATTTTATAGGGGAAGAATG	6791
Db	1923	AGCTTTAGACTTGTAAACCGCTGARAGAGGGGAACTGTTTATTTTATAGGGGAAGAATG	1982
QY	6792	CTGTTATTTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCG	6851
Db	1983	CTGTTATTTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCG	2042
QY	6852	AATCAACGTAAGCAGAGAGCTTCGAAACACTTGGACCTCTGGGCGCTCTCAGCCCATG	6911
Db	2043	AATCAACGTAAGCAGAGAGCTTCGAAACACTTGGACCTCTGGGCGCTCTCAGCCCATG	2102
QY	6912	GATGCCCTGGATTTCTCCCTTTTAGGACCTCTAGCAGCTAATAATTTGCTACTCTCTT	6971
Db	2103	GATGCCCTGGATTTCTCCCTTTTAGGACCTCTAGCAGCTAATAATTTGCTACTCTCTT	2162
QY	6972	TGGACCTGATCTTTTACCTCTTGTAACTTTGTTTAACTTTGTTCTTCCAGAAATCGAAGCTG	7031
Db	2163	TGGACCTGATCTTTTACCTCTTGTAACTTTGTTTAACTTTGTTCTTCCAGAAATCGAAGCTG	2222
QY	7032	ACTCAAAATGAGCCCAAGATGCACTCAAGACTTAAAGATCTACGACAGACCCCTGACCG	7091
Db	2223	ACTCAAAATGAGCCCAAGATGCACTCAAGACTTAAAGATCTACGACAGACCCCTGACCG	2282
QY	7092	GCCTGYTAGCCCAAGATCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAGGAAATCTC	7151
Db	2283	GCCTGYTAGCCCAAGATCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAGGAAATCTC	2342
QY	7152	AGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAAGCTTAGAGGGTSGTGGCC	7211
Db	2343	AGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAAGCTTAGAGGGTSGTGGCC	2402
QY	7212	AACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGACCTGAGAGCAGGACTAG	7271
Db	2403	AACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGACCTGAGAGCAGGACTAG	2462
QY	7272	CTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTTAGSTGGGAGGTGACACATCCAC	7331
Db	2463	CTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTTAGSTGGGAGGTGACACATCCAC	2522
QY	7332	CTTTAAACACGGGGTTCGAACTTAGYTCAACCTGACCAATCAGAGAGCTCACTAAAT	7391
Db	2523	CTTTAAACACGGGGTTCGAACTTAGYTCAACCTGACCAATCAGAGAGCTCACTAAAT	2582
QY	7392	GCTAATTAGGCAAGACAGAGGTAAGAAATAGCCAATCATYATTTGWTGAGAGCACA	7451
Db	2583	GCTAATTAGGCAAGACAGAGGTAAGAAATAGCCAATCATYATTTGWTGAGAGCACA	2642
QY	7452	GCAGGAGGCAATGATCGGGATATAAACCCAAAGTTCGAGCCGGCAACCGCAACCCCC	7511
Db	2643	GCAGGAGGCAATGATCGGGATATAAACCCAAAGTTCGAGCCGGCAACCGCAACCCCC	2702

QY	7512	TTTGGGTCCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTT	7557
Db	2703	TTTGGGTCCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTT	2748
RESULT 2			
CR617248			
LOCUS			
DEFINITION	CR617248	2749 bp mRNA linear HTC 21-JUL-2004	
ACCESSION		full-length cDNA clone CS0D1022YJ18 of Placenta Cot 25-normalized	
VERSION	CR617248	of Homo sapiens (human).	
KEYWORDS	CR617248.1	GI:50498055	
SOURCE	HTC; CNS1 cdNA		
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Homo sapiens		
TITLE	Homo sapiens		
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REMARK	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
	1 (bases 1 to 2749)		
	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
	Full-length cDNA libraries and normalization		
	Unpublished		
	Contact : Feng Liang Email : fliang@lifetech.com URL :		
	http://fulllength.invitrogen.com/ InvitroGen Corporation 1600		
	Paraday Avenue		
REFERENCE	2 (bases 1 to 2749)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :		
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
COMMENT	1st strand cDNA was primed with a NotI-oligo (dr) primer. Five prime		
	end enriched, double-strand cDNA was digested with Not I and cloned		
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library		
	was normalized. Library was constructed by Life Technologies, a		
	division of Invitrogen		
FEATURES			
source	Location/Qualifiers		
	1..2749		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0D1022YJ18"		
	/tissue_type="Placenta Cot 25-normalized"		
	/plasmid="pCMVSPORT_6"		
ORIGIN			
Query Match	29.1%;	Score 2176.4;	DB 6; Length 2749;
Best Local Similarity	98.4%;	Pred. No. 0;	
Matches 2167;	Conservative 34;	Mismatches 0;	Indels 1; Gaps 1;
QY	5352	GATACATCTCTGGGAAGGACCTTACCAGTCATTTTATVACCCCAACTGCGGTTAAAGTG	5411
Db	549	GATACATCTCTGGGAAGGACCTTACCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG	608
QY	5412	GCTGGAGTGAGGTTTGGATACATCACACTTGAGTCAAATCCTGGATCTGCCAAAGGAA	5471
Db	609	GCTGGAGTGAGGTTTGGATACATCACACTTGAGTCAAATCCTGGATCTGCCAAAGGAA	668
QY	5472	CCTGAAATCCAGGAGACAACGCTAGCTATTCTCTGTGAACCTCTAGAGATTGGGCTG	5531
Db	669	CCTGAAATCCAGGAGACAACGCTAGCTATTCTCTGTGAACCTCTAGAGATTGGGCTG	728
QY	5532	CTCTTCAACAACAACGAGGAGAAAGTAACATAAATCATAAATCCCCATGSGCCTCCC	5591
Db	729	CTCTTCAACAACAACGAGGAGAAAGTAACATAAATCATAAATCATAAATCCCCATGSGCCTCCC	787
QY	5592	TTATCATATTTTCTCTKTAATGTTSTTTTACCTSTTTTCACTCTCACTGCACACCCCTCC	5651
Db	788	TTATCATATTTTCTCTKTAATGTTSTTTTACCTCTTTTCACTCTCACTGCACACCCCTCC	847
QY	5652	ATGCCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGATGACAGCTCC	5711

```
QY 6792 CTGTTATTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 6851
Db 1988 CTGTTATTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 2047
QY 6852 AATACAAAGTAKAGCAGAGAGAGCTTCGAAACACTGAGACCTGGGGCCCTCCTCAGCCCATG 6911
Db 2048 AATACAAAGTAKAGCAGAGAGAGCTTCGAAACACTGAGACCTGGGGCCCTCCTCAGCCCATG 2107
QY 6912 GATGCCCTGGATTTCCCTCTTATAGACCTCTAGCAGCTATAATATTGCTACTCTCTCT 6971
Db 2108 GATGCCCTGGATTTCCCTCTTATAGACCTCTAGCAGCTATAATATTGCTACTCTCTCT 2167
QY 6972 TGGACCTGTATCTTTTACCTCTTTTAACTTTTCTCTTCCAGAAATCGAGCTGTAA 7031
Db 2168 TGGACCTGTATCTTTTAACTTTTAACTTTTCTCTTCCAGAAATCGAGCTGTAA 2227
QY 7032 ACTACAAATGGAGCCCAAGATGACAGTCCAAAGACTAAGATCTACCCAGACCCCTGGACCG 7091
Db 2228 ACTACAAATGGAGCCCAAGATGACAGTCCAAAGACTAAGATCTACCCAGACCCCTGGACCG 2287
QY 7092 GCTGYTAGCCCAAGATGATGTTTAAATGACATCAAAAGCACCCCTCTCTGAGGAAATCTC 7151
Db 2288 GCTGTAGCCCAAGATGATGTTTAAATGACATCAAAAGCACCCCTCTCTGAGGAAATCTC 2347
QY 7152 AGCTGACAACTCTACTAGCCGCCAATTCAGCAGAGAGCAGTTAGAGGGGTGTGCGGC 7211
Db 2348 AGCTGACAACTCTACTAGCCGCCAATTCAGCAGAGAGCAGTTAGAGGGGTGTGCGGC 2407
QY 7212 AACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGATGGGAGCTGAGAGCAGGACTAG 7271
Db 2408 AACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGATGGGAGCTGAGAGCAGGACTAG 2467
QY 7272 CTGGATTTCTAGGCTGATTAAGAAATCCYTAAGCCTTAGTGGGAAGGTGACCAATCCAC 7331
Db 2468 CTGGATTTCTAGGCTGATTAAGAAATCCYTAAGCCTTAGTGGGAAGGTGACCAATCCAC 2527
QY 7332 CTTTAAACAGGGGGTTGCACTTAGTGTACACCTGACCAATCAGAGAGCTCACTAAAT 7391
Db 2528 CTTTAAACAGGGGGTTGCACTTAGTGTACACCTGACCAATCAGAGAGCTCACTAAAT 2587
QY 7392 GCTAATTAGGCAAGAGAGAGTAAAGAAATAGCCAATCATYTTATGCTGAGAGCACA 7451
Db 2588 GCTAATTAGGCAAGAGAGAGTAAAGAAATAGCCAATCATYTTATGCTGAGAGCACA 2647
QY 7452 GCAGGAGGACAATGATCGGGATATAAACCCAAAGTTCGAGCCCGCAACGCAACCCCC 7511
Db 2648 GCAGGAGGACAATGATCGGGATATAAACCCAAAGTTCGAGCCCGCAACGCAACCCCC 2707
QY 7512 TTTGGTCCCTCCCTTTGTTATGGAGCTCTGTTTCATGCT 7553
Db 2708 TTTGGTCCCTCCCTTTGTTATGGAGCTCTGTTTCATGCT 2749
```

## RESULT 3

```
CR622175 2500 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized
```

```
DEFINITION full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized
ACCESSION CR622175
VERSION CR622175.1 GI:50502982
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2500)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
```

```
REFERENCE Paraday Avenue
AUTHORS 2 (bases 1 to 2500)
JOURNAL Direct Submission
COMMENT Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES Location/Qualifiers
source 1..2500
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
```

## ORIGIN

```
Query Match 29.0%; Score 2174.4; DB 6; Length 2500;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2165; Conservative 34; Mismatches 0; Indels 1; Gaps 1;
QY 5352 GATACATCCTGGGAAGGACCCTACCCAGTCAATTTATTTATCCCAACTGCGGTTAAAGTG 5411
Db 302 GATACATCCTGGGAAGGACCCTACCCAGTCAATTTATTTATCCCAACTGCGGTTAAAGTG 361
QY 5412 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTGTGATACTGCAAGGAA 5471
Db 362 GCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTGTGATACTGCAAGGAA 421
QY 5472 CTTGAAATCAGAGAGCAACGCTAGCTATTTCTGTGAACTCTAGAGAGTTTGGCGCTG 5531
Db 422 CTTGAAATCAGAGAGCAACGCTAGCTATTTCTGTGAACTCTAGAGAGTTTGGCGCTG 481
QY 5532 CTCTTCAAAACAACAGGAGGAAAGTAACTAAATCATAAATCAATATCCCACTGSCCTCCC 5591
Db 482 CTCTTCAAAACAACAGGAGGAAAGTAACTAAATCATAAATCAATATCCCACTGSCCTCCC 540
QY 5592 TTATCATATTTTCTCTKTASTGTTTSTTTTACCTCTTCTCACTGCAACCCCTCCC 5651
Db 541 TTATCATATTTTCTCTTACTGTTCTTTTACCTCTTCTCACTGCAACCCCTCCC 600
QY 5652 ATGCGCTGTATGACCACTAGCTCCCTTACGAGAGTTTCTATGGAGAGTGCAGCGTCC 5711
Db 601 ATGCGCTGTATGACCACTAGCTCCCTTACGAGAGTTTCTATGGAGAGTGCAGCGTCC 660
QY 5712 CGGAATATTTGATGCCCATCGTATAGAGTCTTTSTAAGGAAACCCCACTTCACTGC 5771
Db 661 CGGAATATTTGATGCCCATCGTATAGAGTCTTTSTAAGGAAACCCCACTTCACTGC 720
QY 5772 CCACACCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCAATCAATAC 5831
Db 721 CCACACCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCAATCAATAC 780
QY 5832 TCATTATTGGACAGGAAATAATGATTAAATCTTAGTTGTCTCTGAGAGCTTGGAGTCACTGT 5891
Db 781 TCATTATTGGACAGGAAATAATGATTAAATCTTAGTTGTCTCTGAGAGCTTGGAGTCACTGT 840
QY 5892 CTGTTGGACTTACTTACCCCAACTGTTATGCTGATGGGGGTGAGTTCAAGATCAGGC 5951
Db 841 CTGTTGGACTTACTTACCCCAACTGTTATGCTGATGGGGGTGAGTTCAAGATCAGGC 900
QY 5952 AAGAGAAAAACATGTAAAGAGAGTAACTCTCCCAACTCAACCCGGGTACATGGCACTCTAG 6011
Db 901 AAGAGAAAAACATGTAAAGAGAGTAACTCTCCCAACTCAACCCGGGTACATGGCACTCTAG 960
QY 6012 CCCCTCAAGAGGACTAGATCTCTCAAACTACATCAAAACCCCTCCGTAACCATCTGCGCT 6071
Db 961 CCCCTCAAGAGGACTAGATCTCTCAAACTACATCAAAACCCCTCCGTAACCATCTGCGCT 1020
```

```
QY 6072 GGTAAAGCCTATTAAATACACACCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTAC 6131
Db 1021 GGTAAAGCCTATTAAATACACACCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTAC 1080
QY 6132 TAACTGTTGGATATGCTCCCTCCCTGAACTTCAAGCCCATATGTTTCAATCCCTGACTGA 6191
Db 1081 TAACTGTTGGATATGCTCCCTCCCTGAACTTCAAGCCCATATGTTTCAATCCCTGACTGA 1140
QY 6192 ACAATGAAACAACCTTCAGCAAGAAATAACACACACTTCGGTTTTAGTAGGACCTCTTGT 6251
Db 1141 ACAATGAAACAACCTTCAGCAAGAAATAACACACACTTCGGTTTTAGTAGGACCTCTTGT 1200
QY 6252 TTCCAATSTGGAATAACCATACCTCAACCTCACTGTGTAAATTTAGCAATACTAC 6311
Db 1201 TTCCAATCTGGAATAACCATACCTCAACCTCACTGTGTAAATTTAGCAATACTAC 1260
QY 6312 ATACACAAACCAACTCCCAATGCACTAGGTGGGTAACTCCCTCCACACAAATAGCTGCGCT 6371
Db 1261 ATACACAAACCAACTCCCAATGCACTAGGTGGGTAACTCCCTCCACACAAATAGCTGCGCT 1320
QY 6372 ACCCTCAGGAATATTTTTGCTGTGCTACCTCAGCCTATCGTTGTTGAATGGCTCTTC 6431
Db 1321 ACCCTCAGGAATATTTTTGCTGTGCTACCTCAGCCTATCGTTGTTGAATGGCTCTTC 1380
QY 6432 AGAATCTATGTGCTCTCTCATCTTAGTGCCCCCCTATGRCATCTACACTGAAACAAGA 6491
Db 1381 AGAATCTATGTGCTCTCTCATCTTAGTGCCCCCCTATGRCATCTACACTGAAACAAGA 1440
QY 6492 TTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTCT 6551
Db 1441 TTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTCT 1500
QY 6552 TATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCAACACCTCTAC 6611
Db 1501 TATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCAACACCTCTAC 1560
QY 6612 TCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTCCGCACTC 6671
Db 1561 TCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTCCGCACTC 1620
QY 6672 CTTGGTCACTTCGCAAGTCACTTAACCTCCCTAGCAGCAGTAGTCTTCRAAATCGAAG 6731
Db 1621 CTTGGTCACTTCGCAAGTCACTTAACCTCCCTAGCAGCAGTAGTCTTCRAAATCGAAG 1680
QY 6732 AGCTTTAGACTTGCCTAACCGCTGARAGAGGGGAACCTGTTTATTTTGGGGAAGAAATG 6791
Db 1681 AGCTTTAGACTTGCCTAACCGCTGARAGAGGGGAACCTGTTTATTTTGGGGAAGAAATG 1740
QY 6792 CTGTTATTAATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 6851
Db 1741 CTGTTATTAATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 1800
QY 6852 AATACAACTAGAGAGAGAGGCTTGAACACTGACCCCTGGGGCTCTCAGCCCATG 6911
Db 1801 AATACAACTAGAGAGAGAGGCTTGAACACTGACCCCTGGGGCTCTCAGCCCATG 1860
QY 6912 GATGCCCTGATTTCTCCCTCTTAGGACTCTAGCAGCTATAATATGCTACTCTCTCT 6971
Db 1861 GATGCCCTGATTTCTCCCTCTTAGGACTCTAGCAGCTATAATATGCTACTCTCTCT 1920
QY 6972 TGGACCTGATCTTTTACCTCTGTTAACTTTGCTCTTCAGAAATCGAAGCTGTAA 7031
Db 1921 TGGACCTGATCTTTTACCTCTGTTAACTTTGCTCTTCAGAAATCGAAGCTGTAA 1980
QY 7032 ACTACAAATGGAGCCCAAGATGAGTCCAGACTAAGATCTACCGCAGACCCCTGAGACG 7091
Db 1981 ACTACAAATGGAGCCCAAGATGAGTCCAGACTAAGATCTACCGCAGACCCCTGAGACG 2040
QY 7092 GCTGTGTAGCCACGATCTGATGTTAATGACATCAAGGACCCCTCTGAGGAATCTC 7151
Db 2041 GCTGTGTAGCCACGATCTGATGTTAATGACATCAAGGACCCCTCTGAGGAATCTC 2100
```

```
QY 7152 AGTGCACACACCTCTACTAGCCGCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 7211
Db 2101 AGTGCACACACCTCTACTAGCCGCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTCGGCC 2160
QY 7212 AACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACTCAGAGAGAGGACTAG 7271
Db 2161 AACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACTCAGAGAGAGGACTAG 2220
QY 7272 CTGGATTTCTAGGCTGAYTAAGAATCCYTAAGCCTAGTAGGGAAGGTGACCAATCCAC 7331
Db 2221 CTGGATTTCTAGGCTGAYTAAGAATCCYTAAGCCTAGTAGGGAAGGTGACCAATCCAC 2280
QY 7332 CTTTAAACAGGGGCTTGCAACTTAGTGTACACCTGACCAATCAGAGAGCTCACTAAAT 7391
Db 2281 CTTTAAACAGGGGCTTGCAACTTAGTGTACACCTGACCAATCAGAGAGCTCACTAAAT 2340
QY 7392 GCTAATTAGCAAGACAGGAGTAAAGAAATAGCAATCATVTATTGWTGAGAGACACA 7451
Db 2341 GCTAATTAGCAAGACAGGAGTAAAGAAATAGCAATCATCTATTGCTGAGAGACACA 2400
QY 7452 GCAGGAGGACCAATGATCGGGATATAAACCCCAAGTTCGAGCCGCAACGGCAACCCCC 7511
Db 2401 GCAGGAGGACCAATGATCGGGATATAAACCCCAAGTTCGAGCCGCAACGGCAACCCCC 2460
QY 7512 TTTGGTCCCTCCCTTTGTTATGGGAGCTCTGTTTTCATG 7551
Db 2461 TTTGGTCCCTCCCTTTGTTATGGGAGCTCTGTTTTCATG 2500
```

## RESULT 4

CR625046

LOCUS

DEFINITION

2748 bp mRNA linear HTC 21-JUL-2004

full-length cDNA clone CS0D1044YK06 of Placenta Cot 25-normalized

of Homo sapiens (human)

ACCESSION

CR625046

VERSION

HTC; CNSLT cDNA

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 2748)

AUTHORS

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue

Genoscope.

REFERENCE

2 (bases 1 to 2748)

AUTHORS

Direct Submission

JOURNAL

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

FEATURES

Location/Qualifiers

1..2748

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1044YK06"

/tissue type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT\_6"

ORIGIN

Query Match 29.0%; Score 2169.4; DB 6; Length 2748;

Best Local Similarity 98.4%; Pred. No. 0;

Matches 2160; Conservative 34; Mismatches 0; Indels 1; Gaps 1;

QY	5352	GATACATCTGGGAGGACCCCTACCCAGCTCATTTTATYTAACCCAACTGCGGTTAAAGTG	5411
Db	555	GATACATCTGGGAGGACCCCTACCCAGCTCATTTTATCTACCCAACTGCGGTTAAAGTG	614
QY	5412	GCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAAAATCTGGATACTGCAAGGAA	5471
Db	615	GCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAAAATCTGGATACTGCAAGGAA	674
QY	5472	CCTGAAATCCAGGAGACAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATTTGGCGCTG	5531
Db	675	CCTGAAATCCAGGAGACAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATTTGGCGCTG	734
QY	5532	CTCTTCAACAACAACGAGGAGAAAGTAACTAAATCATAAATCCCCCATGGSCCTCCC	5591
Db	735	CTCTTCAACAACAACGAGGAGAAAGTAACTAAATCATAAATCCCCCATGGSCCTCCC	793
QY	5592	TTATCATATTTTCTCTASTGTSTTTTAYACCTSTTTTCACTCTCACTGCAACCCCTCC	5651
Db	794	TTATCATATTTTCTCTTACTGTCTTTTACCCCTCTTTCACTCTCACTGCAACCCCTCC	853
QY	5652	ATGCGCGTGTATGACCAAGTAGCTCCCTTACCMAGAGTTTCTATGGAGAAATGCAAGCTCC	5711
Db	854	ATGCGCGTGTATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGGAGAAATGCAAGCTCC	913
QY	5712	CGGAAATATGTAGTCCCATCTGTATAGAGTCTTTTAAAGGAAACCCCACTTCACTGC	5771
Db	914	CGGAAATATGTAGTCCCATCTGTATAGAGTCTTTTAAAGGAAACCCCACTTCACTGC	973
QY	5772	CCACACCATATGCCCGCACTGTATCACTCTGCCACTCTTTCATGCAATGCATGCAAAATAC	5831
Db	974	CCACACCATATGCCCGCACTGTATCACTCTGCCACTCTTTCATGCAATGCATGCAAAATAC	1033
QY	5832	TCATTTATGGACAGGAAATGATTAATCTTAGTTGTCTGGAGGACTTGGAGTCACTGT	5891
Db	1034	TCATTTATGGACAGGAAATGATTAATCTTAGTTGTCTGGAGGACTTGGAGTCACTGT	1093
QY	5892	CTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC	5951
Db	1094	CTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGC	1153
QY	5952	AGAGAAACATGTAAAGAGTAACTCCCAACTCACCSGGGTACATGGCACCTCTAG	6011
Db	1154	AGAGAAACATGTAAAGAGTAACTCCCAACTCACCSGGGTACATGGCACCTCTAG	1213
QY	6012	CCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTTCCGTTACCCATCTCGCT	6071
Db	1214	CCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTTCCGTTACCCATCTCGCT	1273
QY	6072	GGTAAGCCTATTTAATACACCTCACTGGGCTTCAATGAGTCTCGGCCCAAAACCTTAC	6131
Db	1274	GGTAAGCCTATTTAATACACCTCACTGGGCTTCAATGAGTCTCGGCCCAAAACCTTAC	1333
QY	6132	TAACTGTTGGATATGCTCCCTGAACTTTCARGCCATATGTTCAATCCCTGTACCTGA	6191
Db	1334	TAACTGTTGGATATGCTCCCTGAACTTTCARGCCATATGTTCAATCCCTGTACCTGA	1393
QY	6192	ACAAATGGAAACAATTCAGCACAGAAATAAACACCACTTCCGTTTATGAGGACCTCTGT	6251
Db	1394	ACAAATGGAAACAATTCAGCACAGAAATAAACACCACTTCCGTTTATGAGGACCTCTGT	1453
QY	6252	TTTCAATSTGGAAATAACCACTCACTCAAACTCACTGTTGTAATAATTTAGCAATCTAC	6311
Db	1454	TTTCAATSTGGAAATAACCACTCACTCAAACTCACTGTTGTAATAATTTAGCAATCTAC	1513
QY	6312	ATACACAACTCCCAATGATCAGGTGGGTAACTCTCCCAACCAATAGTCTGCT	6371
Db	1514	ATACACAACTCCCAATGATCAGGTGGGTAACTCTCCCAACCAATAGTCTGCT	1573
QY	6372	ACCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAAATGGCTCTTC	6431
Db	1574	ACCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAAATGGCTCTTC	1633

QY	6432	AGAATCTATGTGCTTCTCTCATTTAGTCCCCCVATGRCCATCTACACTGAACAAGA	6491
Db	1634	AGAATCTATGTGCTTCTCTCATTTAGTCCCCCVATGRCCATCTACACTGAACAAGA	1693
QY	6492	TTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATCTTCTCTTTGT	6551
Db	1694	TTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATCTTCTCTTTGT	1753
QY	6552	TATAGGAGCAGGAGTGTAGTGCACTAGGTACTGGGATTTGGCGGTATCAAACTCTAC	6611
Db	1754	TATAGGAGCAGGAGTGTAGTGCACTAGGTACTGGGATTTGGCGGTATCAAACTCTAC	1813
QY	6612	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCCGCACTC	6671
Db	1814	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCCGCACTC	1873
QY	6672	CCTGGTCACTTGCAGAGATCAACTTAATCTCCTAGCAGCAGTAGTCTTTCRAAATCGAAG	6731
Db	1874	CCTGGTCACTTGCAGAGATCAACTTAATCTCCTAGCAGCAGTAGTCTTTCRAAATCGAAG	1933
QY	6732	AGCTTTAGACTTGTCTAACCGCTGARAGAGGGGAACTGTGTTTATTTTAGGGGAAGAATG	6791
Db	1934	AGCTTTAGACTTGTCTAACCGCTGARAGAGGGGAACTGTGTTTATTTTAGGGGAAGAATG	1993
QY	6792	CTGTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCGAGATCG	6851
Db	1994	CTGTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCGAGATCG	2053
QY	6852	AATACAACTAGACAGAGAGCTTTCGAAAACACTGGGACCTTGGGGCCCTCTCAGCCRAATG	6911
Db	2054	AATACAACTAGACAGAGAGCTTTCGAAAACACTGGGACCTTGGGGCCCTCTCAGCCRAATG	2113
QY	6912	GATGCCCTGGATTTCTCCCTTTTAGGACCTCTAGCAGCTATAATATTTGCTACTCTCTT	6971
Db	2114	GATGCCCTGGATTTCTCCCTTTTAGGACCTCTAGCAGCTATAATATTTGCTACTCTCTT	2173
QY	6972	TGGACCTCTGATCTTTTACCTCTTGTAACTTTGCTCTTCCAGAAATCGAGGCTGTAA	7031
Db	2174	TGGACCTCTGATCTTTTACCTCTTGTAACTTTGCTCTTCCAGAAATCGAGGCTGTAA	2233
QY	7032	ACTACAAATGAGGCCCAAGATGCACTCAAGACTTAAGACTTACCGCAGACCCCTCGGACCG	7091
Db	2234	ACTACAAATGAGGCCCAAGATGCACTCAAGACTTAAGACTTACCGCAGACCCCTCGGACCG	2293
QY	7092	GCCTGTAGCCCAAGATCTGATGTTAATGAATCAAAAGGACCCCTCTCTGAGGAAATCTC	7151
Db	2294	GCCTGTAGCCCAAGATCTGATGTTAATGAATCAAAAGGACCCCTCTCTGAGGAAATCTC	2353
QY	7152	AGCTGCAACACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGGGTSGTCCGCC	7211
Db	2354	AGCTGCAACACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGGGTSGTCCGCC	2413
QY	7212	AACCTCCCAACAGCAGCTTAGTGTTCCTGTTGAGATGGGGAATGAGAGACAGGACTAG	7271
Db	2414	AACCTCCCAACAGCAGCTTAGTGTTCCTGTTGAGATGGGGAATGAGAGACAGGACTAG	2473
QY	7272	CTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGAACACATCCAC	7331
Db	2474	CTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTGAACACATCCAC	2533
QY	7332	CTTTAAACACCGGGCTTGCAACTTAGYTTCACACCTGACCAATTCAGAGAGCTCACTAAAT	7391
Db	2534	CTTTAAACACCGGGCTTGCAACTTAGYTTCACACCTGACCAATTCAGAGAGCTCACTAAAT	2593
QY	7392	GCTAATTAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTTGWTGAGAGACA	7451
Db	2594	GCTAATTAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTTGWTGAGAGACA	2653
QY	7452	GCAGGAGGACAATGATCGGATATAAACCCAAAGTTCGAGCCCGGCAACCGGCAACCC	7511
Db	2654	GCAGGAGGACAATGATCGGATATAAACCCAAAGTTCGAGCCCGGCAACCGGCAACCC	2713
QY	7512	TTTGGGTCCCTCCCTTTGTTGATGGGAGCTCTGTTT	7546





```

1984  CTTGTTATTATGTTAAATCAATCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2043
QY 6852  AATAACAAGTAKAGCAGAGAGCTTCGAAACACTGAGACCTGGGCGCTCCTCAGCCRAATG 6911
Db 2044  AATAACAAGTAKAGCAGAGAGCTTCGAAACACTGAGACCTGGGCGCTCCTCAGCCRAATG 2103
QY 6912  GATGCCCTGGATCTCCCTCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCTCTT 6971
Db 2104  GATGCCCTGGATCTCCCTCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCTCTT 2163
QY 6972  TGGACCTCTGATCTTTTACCTCTCTTCTTAACTTCTCTTCCAGAAATCGAAGCTGTAA 7031
Db 2164  TGGACCTCTGATCTTTTACCTCTCTTCTTAACTTCTCTTCCAGAAATCGAAGCTGTAA 2223
QY 7032  ACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGACTTACCGCAGACCCCTGGACCG 7091
Db 2224  ACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGACTTACCGCAGACCCCTGGACCG 2283
QY 7092  GCTGTTAGCCCAAGATCTGATGTTAATGACATCAAGGACCCCTCTAGGAAATCTC 7151
Db 2284  GCTGTTAGCCCAAGATCTGATGTTAATGACATCAAGGACCCCTCTAGGAAATCTC 2343
QY 7152  ACCTGCACAACTCTACTACGCCCAATTTACGACGAGACGTTAGAGCGTSGTCGCC 7211
Db 2344  ACCTGCACAACTCTACTACGCCCAATTTACGACGAGACGTTAGAGCGTSGTCGCC 2403
QY 7212  AACCTCCCCAACAGCAGCTTAGGTTTTCTGTTGAGATGGGGACTGAGAGACGAGCTAG 7271
Db 2404  AACCTCCCCAACAGCAGCTTAGGTTTTCTGTTGAGATGGGGACTGAGAGACGAGCTAG 2463
QY 7272  CTGGATTTCTTAGCTGATGATTAAGATTCCTTAAGCTTAGTGGGAGGTGACACATCCAC 7331
Db 2464  CTGGATTTCTTAGCTGATGATTAAGATTCCTTAAGCTTAGTGGGAGGTGACACATCCAC 2523
QY 7332  CTTTAAACAGCGGGCTTCACTTAGTCAACCTGACCAATGACAGAGCTCACTAAAT 7391
Db 2524  CTTTAAACAGCGGGCTTCACTTAGTCAACCTGACCAATGACAGAGCTCACTAAAT 2583
QY 7392  GCTAATTAGGCAAGACAGAGCTAAAGAAATAGCAATCATYTTATTGCTGAGAGCACA 7451
Db 2584  GCTAATTAGGCAAGACAGAGCTAAAGAAATAGCAATCATYTTATTGCTGAGAGCACA 2643
QY 7452  GCAGAGGACAAATGATCGGATATAAACCCAAAGTTCGAGCGGCAACGGCAACCCCC 7511
Db 2644  GCAGAGGACAAATGATCGGATATAAACCCAAAGTTCGAGCGGCAACGGCAACCCCC 2703
QY 7512  TTTGGGTCCCTC 7524
Db 2704  TTTGGGTCCCTC 2716

```

## RESULT 6

LOCUS BX365066/c 1071 bp mRNA linear EST 08-APR-2004  
DEFINITION BX365066 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1046YA18 3-PRIME, mRNA sequence.

ACCESSION BX365066

VERSION BX365066.2 GI:46304105

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.

## REFERENCE

1 (bases 1 to 1071)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 5, 2003 this sequence version replaced gi:30374869.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4215.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS1A1012ZC10NP1&c=4215.f.r.

## FEATURES

source

1. .1071

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="CS0D1046YA18"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

```

Query Match      12.9%; Score 969; DB 4; Length 1071;
Best Local Similarity 95.8%; Pred. No. 1.2e-272;
Matches 985; Conservative 25; Mismatches 16; Indels 2; Gaps 2;

QY 6520  CGCAACAAAGAGTACCCATTCTTCCTTTTGTATAGGAGCAGGAGTCTAGTGCACTA 6579
Db 1026  CCCSCACAAAGAGWCCCATVCTCCYTTTATAGGAGCAGGAGTGCTA-GTGCACTA 968
QY 6580  GGTACTGCAATGGCGGTATCACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAA 6639
Db 967  GGGACTGCAATGGCGGTATACCAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAA 908
QY 6640  CTAATATGGGACATGGAACGGTGCAGCTCCCTGGTCACTTCCAAAGATCAACTTAAC 6699
Db 907  CTAATATGGGACATGGAACGGTGCAGCTCCCTGGTCACTTCCAAAGATCAACTTAAC 848
QY 6700  TCCCTAGCAGCAGTAGTCTTTCRAAATCGAAGAGCTTTAGACTYCTTAAACCGCTGAGA 6759
Db 847  TCCCTAGCAGCAGTAGTCTTTCRAAATCGAAGAGCTTTAGACTYCTTAAACCGCTGAGA 788
QY 6760  GGGGGAACCTGTTTATTTTATAGGGAAGAAATGCTGTTTATTTATTTATTTATTTAT 6819
Db 787  GGGGGAACCTGTTTATTTTATAGGGAAGAAATGCTGTTTATTTATTTATTTATTTAT 728
QY 6820  GTCACTGAGAAAGTTTAAAGAAATTCGAGATCGAATCAACACGTAACAGCAGAGAGCTT 6879
Db 727  GTCACTGAGAAAGTTTAAAGAAATTCGAGATCGAATCAACACGTAACAGCAGAGAGCTT 668
QY 6880  AACACTGACCCCTGGGCGCTCTCAGCCRAATGAGTGGCTGAGTCTTCCCTCTTTAGGA 6939
Db 667  AACACTGACCCCTGGGCGCTCTCAGCCRAATGAGTGGCTGAGTCTTCCCTCTTTAGGA 608
QY 6940  CCTCTAGCAGCTATAATATTGCTACTCTCTTTGACCCCTGATCTTTTAACTCTCTGTT 6999
Db 607  CCTCTAGCAGCTATAATATTGCTACTCTCTTTGACCCCTGATCTTTTAACTCTCTGTT 548
QY 7000  AACTTTGTCTCTTCCAGAAATCGAAGCTGTAAACTTACAAATCGAGCCCAAGATGCACT 7059
Db 547  AACTTTGTCTCTTCCAGAAATCGAAGCTGTAAACTTACAAATCGAGCCCAAGATGCACT 488
QY 7060  AAGACTAAGATCTACCGCAGACCCCTGAGCCGGCTGTYTAGCCACGATCTGATGTTAAT 7119
Db 487  AAGACTAAGATCTACCGCAGACCCCTGAGCCGGCTGTYTAGCCACGATCTGATGTTAAT 428
QY 7120  GACATCAAGGACCCCTCTCAGGAAATCTCAGCTGCAACACCTCTACTAGCCCAAT 7179
Db 427  GACATCAAGGACCCCTCTCAGGAAATCTCAGCTGCAACACCTCTACTAGCCCAAT 368
QY 7180  TCAGCAGGAAGCAGTTAGAGCGGTGTCGGCCAACTCCGCCAACAGCAGCTTAGGTTTCC 7239

```

Db 367 TCAGCAGGAAGCAGTTAGAGGGTCTGTCGGCCAACTCCCAACAGCACTTAGGTTTCC 308

QY 7240 TGTTCAGATCGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCC 7299

Db 307 TGTTCAGATCGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCC 248

QY 7300 YTAAGCCTAGTCTGGGAAGGTGACACATCCACCTTTAAACACGGGGCTTCAACTTAGYT 7359

Db 247 CTAAAGCTAGCTGGGAAGGTGACACATCCACCTTTAAACACGGGGCTTCAACTTAGCT 188

QY 7360 CACACCTGACCAATACAGAGAGCTCACTAAATGCTAAATTTAGGCAAGACAGAGGTAAG 7419

Db 187 CACACCTGACCAATACAGAGAGCTCACTAAATGCTAAATTTAGGCAAGACAGAGGTAAG 128

QY 7420 AATAGCCATCATYATTTCMTGAGAGCAGACAGAGGAGCAATGATCGGATATAA 7479

Db 127 AATAGCCATCATCTATTGCTTGGAGCAGACAGAGGAGCAATGATCGGATATAA 68

QY 7480 CCCAAGTTCGAGCGCGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGC 7539

Db 67 CCCAAGTTCGAGCGCGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGC 9

QY 7540 TCTGTTTT 7547

Db 8 TCKGTTTT 1

## RESULT 7

BX337769/c

LOCUS

DEFINITION

BX337769

ACCESSION

BX337769

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 998)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 2, 2003 this sequence version replaced gi:30337641.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI051Y1M13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

```
LOCUS BX378303 1058 bp mRNA linear EST 23-APR-2004
DEFINITION BX378303 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1018YH07 3-PRIME, mRNA sequence.
ACCESSION BX378303
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
REFERENCE 1 (bases 1 to 1058)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30439129.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1018CD04NP1&c=4215.r.
FEATURES
Location/Qualifiers
1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1018YH07"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 12.1%; Score 909.6; DB 4; Length 1058;
Best Local Similarity 90.6%; Pred. No. 4e-255;
Matches 949; Conservative 38; Mismatches 56; Indels 4; Gaps 3;
QY 6501 TTATGTCATATCTAAGCCCGCAACAAAGAGTACCATTCTTCCTTTTGTATAGGAGC 6560
DB 1047 TTAWAATAATGYAWATCAASCCSCAAMAAPATAACCATCTCTCTTTKTATAG-AGC 989
QY 6561 AGGAGTGCTAGTGCACTAGTACTGGCAATGCGGATACACAACTCTACTCAGTTCTA 6620
DB 988 AGGAGAGCTAGK--GCATAGTACTGCAATTGGC-GTATCAACAACCTCTACTCAGTTCTA 932
QY 6621 CTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGGACTCCCTGCTAC 6680
DB 931 CTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGGACTCCCTGCTAC 872
QY 6681 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTACCTTCRAAATCGAAGAGCTTTAGA 6740
DB 871 CTTGCAAGATCAACTTAACTCCCTAGCAGCAGTACCTTCRAAATCGAAGAGCTTTAGA 812
QY 6741 CTYGTCTAACCGCTGARAGAGGGGAACTGTTTATTTTATAGGGGAAGAATGCTGTTATTA 6800
DB 811 CTTGCTAACCGCTGAAAGAGGKKWCYTTTTTTTTTTTATAGGGGAAGAATGCTGTTACTA 752
QY 6801 TGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAGAAATTCAGATCGAATCAACG 6860
DB 751 TGTTAATCAATCCGGAATCTCACTGAGAAGTTAAGAAATTCAGATCGAATCAACG 692
QY 6861 TAKAGCAGARGAGCTCTTGAACAACACTGGACCTCGGCGCTCTCAGCGCATGATGCCCTG 6920
DB 691 TAGAGCAGAGGAGCTTCGAAACACTGGACCTCGGCGCTCTCAGCGCATGATGCCCTG 632
```

```
QY 6921 GATTCTCCCTCTTAGGACCTCTAGCAGCTATATATTTGCTACTCTCTTTGGACCTTG 6980
DB 631 GATTCTCCCTCTTAGGACCTCTAGCAGCTATATATTTGCTACTCTCTTTGGACCTTG 572
QY 6981 TATCTTTTACCTCTCTTTTAACTTTGCTCTTTCAGAAATCGAAGCTGTAAATACAAAT 7040
DB 571 TATCTTTTAACTCTCTTTTAACTTTGCTCTTTCAGAAATCGAAGCTGTAAATACAAAT 512
QY 7041 GGAGCCCAAGATGAGTCCAAAGACTAAGATCTACGCGAGACCCCTTGGAGCCGCTGYTAG 7100
DB 511 GGAGCCCAAGATGAGTCCAAAGACTAAGATCTACGCGAGACCCCTTGGAGCCGCTGYTAG 452
QY 7101 CCCAGATCTGATGTTAATGATCAATCAAGGACCCCTCTGAGGAAATCTGAGCTGCACA 7160
DB 451 CCCAGATCTGATGTTAATGATCAATCAAGGACCCCTCTGAGGAAATCTGAGCTGCACA 392
QY 7161 ACCTCTACTACGCCCCCAATTCAGCAGGAGAGTGTAGAGCGGTGTCGCGCAACCTCCCC 7220
DB 391 ACCTCTACTACGCCCCCAATTCAGCAGGAGAGTGTAGAGCGGTGTCGCGCAACCTCCCC 332
QY 7221 AACAGCACTTAGTCTTCTCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGAATTC 7280
DB 331 AACAGCACTTAGTCTTCTCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGAATTC 272
QY 7281 CTAGCTGATTAAGNATCCYTAAGCTAGTAGTGGGAGGTGACCATCCACCTTTTAAACA 7340
DB 271 YTAGCTGACTAAGAAATCCCTAAGCTAGCTGGGAGGTGACCATCCACCTTTTAAACA 212
QY 7341 CGGGCTTTCGAATTTAGTTCACACCTGACCAATCAGAGAGGTCTCAATAATCTAATTAG 7400
DB 211 CGGGCTTTCGAATTTAGTTCACACCTGACCAATCAGAGAGGTCTCAATAATCTAATTAG 152
QY 7401 GCAAGACAGGAGGTAAAGAAATAGCCAAATCATYTTATTGCMWTGAGACACAGCAGAGGG 7460
DB 151 GCAAGACAGGAGGTAAAGAAATAGCCAAATCATCTATTGCTCTGAGACACAGCAGAGGG 92
QY 7461 ACAATGATCGGATATAAACCCCAAGTTCGAGCCGCAACGGCAACCCCTTTGGTCTC 7520
DB 91 ACAATGATCGGATATAAACCCCAAGTTCGAGCCGCAACGGCAACCCCTTTGGTCTC 32
QY 7521 CCTCCCTTTGTATGGGAGCTCTGTTT 7547
DB 31 CMTGCGGCTTGTGTGGAGTTTGT 5
RESULT 9
BX391741/c
LOCUS BX391741 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1015YMI3 3-PRIME, mRNA sequence.
ACCESSION BX391741
VERSION BX391741.1 GI:30619529
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
```

For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?s=CS0BA1022ZA10\\_CS02089\\_1&c=4215.r](http://www.genoscope.cns.fr/cdna?s=CS0BA1022ZA10_CS02089_1&c=4215.r)

FEATURES		Location/Qualifiers	
source		1..935	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0D1051YM13"	
		/tissue_type="PLACENTA COT 25-NORMALIZED"	
		/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"	
		/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN			
Query Match		12.1%; Score 908.6; DB 4; Length 935;	
Best Local Similarity		98.4%; Pred. No. 7.3e-255;	
Matches 919; Conservative 6; Mismatches 8; Indels 1; Gaps 1;			
QY	5805	TGCACCTCTTTCATGATGATCAATATCTATTTGGACAGGAAAAATGATTATCTAG	5864
DB	933	TGCCCCCTTTTGCATGATGATCAATATCTATTTGGACAGGAAAAATGATTATCTAG	875
QY	5865	TTGTCTCGGAGGACTTGGAGTCACTGCTGTGGACTTACTTACCCAAAATGGTATGTC	5924
DB	874	TTGTCTCGGAGGACTTGGAGTCACTGCTGTGGACTTACTTACCCAAAATGGTATGTC	815
QY	5925	TGATGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTCTCCCA	5984
DB	814	TGATGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTCTCCCA	755
QY	5985	ACTCACCGGGTACATGGCACTCTAGCCCTTCAAAAGGACTAGATCTCTCAAAACTACA	6044
DB	754	ACTCACCGGGTACATGGCACTCTAGCCCTTCAAAAGGACTAGATCTCTCAAAACTACA	695
QY	6045	TGAACCCCTCGTACCATCTGCTGGTAAGCTTATTTAATACACCTCTCTAGGGCT	6104
DB	694	TGAACCCCTCGTACCATCTGCTGGTAAGCTTATTTAATACACCTCTCTAGGGCT	635
QY	6105	CCATGAGTCTCGGCCCAAAACCTTACTACTGTGGATGCTCCCTCGTGAACCTTCA	6164
DB	634	CCATGAGTCTCGGCCCAAAACCTTACTACTGTGGATGCTCCCTCGTGAACCTTCA	575
QY	6165	GCCATATGTTTCAATCCCTGTACTGAAATGGAACAACTTCAGCACAGAAATAAACAC	6224
DB	574	GCCATATGTTTCAATCCCTGTACTGAAATGGAACAACTTCAGCACAGAAATAAACAC	515
QY	6225	CACCTCGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT	6284
DB	514	CACCTCGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACT	455
QY	6285	CACCTGTGTAATTTAGCAATCTACTACATACAAACCACTCCCAATGATCAGGTGGT	6344
DB	454	CACCTGTGTAATTTAGCAATCTACTACATACAAACCACTCCCAATGATCAGGTGGT	395
QY	6345	AACTCTCTCCCAACAAATAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTTACCTC	6404
DB	394	AACTCTCTCCCAACAAATAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTTACCTC	335
QY	6405	AGCTATTCGTTTGTGATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTCC	6464
DB	334	AGCTATTCGTTTGTGATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTCC	275
QY	6465	CCCTATGRCCTCTACACTGAAACAGATTTATACAGTTATGTATCTTAAGCCCGCAA	6524
DB	274	CCCTATGACCATCTACACTGAAACAGATTTATACAGTTATGTATCTTAAGCCCGCAA	215
QY	6525	CAAAAGAGTACCCATCTTCTTTTGTATAGAGCAGGAGTCTAGGTGCACTAGGTAC	6584
DB	214	CAAAAGAGTACCCATCTTCTTTTGTATAGAGCAGGAGTCTAGGTGCACTAGGTAC	155

QY	5585	TGCATTTGGCGGTATCACAACTCTTACTAGTTCTTACTACAACTATCTCAAGAACTAAA	6644
DB	154	TGCATTTGGCGGTATCACAACTCTTACTAGTTCTTACTACAACTATCTCAAGAACTAAA	95
QY	6645	TGGGACATGGAACGGGTGCGGACTCCCTGGTCCACCTTGCAGATCAACTTAACTCCCT	6704
DB	94	TGGGACATGGAACGGGTGCGGACTCCCTGGTCCACCTTGCAGATCAACTTAACTCCCT	35
QY	6705	AGCAGCAGTATGCTCTTCRAAATCGAAGAGCTTTA	6738
DB	34	AGCAGCAGTATGCTCTTCRAAATCGAAGAGCTTTA	1
RESULT 10			
BX347952			
LOCUS			
DEFINITION			
BX347952 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
clone CS0D1051YM13 5-PRIME, mRNA sequence.			
ACCESSION			
BX347952			
VERSION			
BX347952.1 GI:30375235			
KEYWORDS			
EST.			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominiidae; Homo.			
1 (bases 1 to 955)			
REFERENCE			
AUTHORS			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE			
Full-length cDNA libraries and normalization			
JOURNAL			
Unpublished (2001)			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime			
end enriched, double-strand cDNA was digested with Not I and cloned			
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library			
was normalized. Library was constructed by Life Technologies, a			
division of Invitrogen. This sequence belongs to sequence cluster			
4215.r			
For more information about this cluster, see			
<a href="http://www.genoscope.cns.fr/cdna?s=CS0BAE023ZF07_AE02167_1&amp;c=4215.r">http://www.genoscope.cns.fr/cdna?s=CS0BAE023ZF07_AE02167_1&amp;c=4215.r</a>			
FEATURES		Location/Qualifiers	
source		1..955	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0D1051YM13"	
		/tissue_type="PLACENTA COT 25-NORMALIZED"	
		/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"	
		/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN			
Query Match		11.8%; Score 885.2; DB 4; Length 955;	
Best Local Similarity		97.4%; Pred. No. 5.8e-248;	
Matches 931; Conservative 8; Mismatches 13; Indels 4; Gaps 4;			
QY	5806	GCCACTCTTTTGCATGATGCAATATCTCATTATTGGACAGGAAAAATGATTATCTAGT	5865
DB	1	GCCACTCTTTTGCATGATGCAATATCTCATTATTGGACAGGAAAAATGATTATCTAGT	60
QY	5866	TGTCCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAAAATGGTATGCT	5925
DB	61	TGTCCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAAAATGGTATGCT	120
QY	5926	GATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTCTCCCA	5985
DB	121	GATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAACTCTCCCA	180



```
QY 7240 TGTGAGTGGGGAGTGGAGAGCAGGACTAGCTGGATTTCTAGGCTGATTAAGATCC 7299
Db 315 TGTGAGTGGGGAGTGGAGAGCAGGACTAGCTGGATYCCYAGGCTGAAGATCC 256
QY 7300 YTAAGCTAGTGGGAAGTGGACACATCCACCTTTAAACACACGGGGCTTGCAACTTAGYT 7359
Db 255 CTAAGCTAGTGGGAAGTGGACACATCCACCTTTAAACACACGGGGCTTGCAACTTAGCT 196
QY 7360 CACACCTGACCAATCAGAGAGCTCACTAAATGCTTAATAGGCAAGAGAGGTTAAG 7419
Db 195 CACACCTGACCAATCAGAGAGCTCACTAAATGCTTAATAGGCAAGAGAGGTTAAG 136
QY 7420 AAATAGCACAATCTATTTATTCCTGAGAGCAGCAGGAGGCAATGATCGGATATAA 7479
Db 135 TAATAKNVRATTATTTATTCCTGAGAGCAGCAGGAGGAGCAATATCGGATATAA 76
QY 7480 CCAAGTTCGAGCGGCAACCGCAACCCCTTTGGTCCCTCCCTTTGTATGGAGC 7539
Db 75 YCCAAGTTCGAGCGGCAACCGCAACCCCTTTGGTCCCTCCCTTTGTATGGAGC 17
QY 7540 TCTGTTTTCATGCTAT 7555
Db 16 TTTTNTTNTTATTT 1
```

## RESULT 12

LOCUS  
DEFINITION  
BX378304 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1018YH07 5-PRIME, mRNA sequence.

ACCESSION  
BX378304

VERSION  
BX378304.2 GI:46569716

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

AUTHORS  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE  
Full-length cDNA libraries and normalization

JOURNAL  
Unpublished (2001)

COMMENT  
On May 8, 2003 this sequence version replaced gi:30443114.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with NotI and cloned

into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS0D1018CD040P1&c=4215.r.

Location/Qualifiers

1..1014  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1018YH07"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

## FEATURES

source

## ORIGIN

Query Match 11.6%; Score 866.6; DB 4; Length 1014;  
Best Local Similarity 95.0%; Pred. No. 1.8e-242;  
Matches 874; Conservative 24; Mismatches 21; Indels 1; Gaps 1;

RESULT 13  
CR749521

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

CR749521 6400 bp mRNA linear HTC 19-AUG-2004  
Homo sapiens mRNA; cDNA DKFP686121167 (from clone DKFP686121167).  
CR749521  
HTC.  
Homo sapiens (human)

980 AAGGGTAAACAGTGAACCAA 999  
990 AAGGGTAAACAGTGAACCAA 919

920 TCCTTCCCACTAATAAGGAGCCCTTTCAACCCCAATATGTCCTTCAAGAGATGACAA 979  
840 TCCTTCCCACTAATAAGGAGCCCTTTCAACCCCAATATGTCCTTCAAGAGATGACAA 899

860 TGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAAGCATCCCGAC 919  
780 TGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCAAGCATCCCGAC 839

800 GAAGTGCCTAAGTACAACTTTCTTTTCAATTAAGAGACAACCTCACTCAATTTATGTAAGAG 859  
720 GAAGTGCCTAAGTACAACTTTCTTTTCAATTAAGAGACAACCTCACTCAATTTATGTAAGAG 779

680 AGACACTAAGAAAGAAACGACTTATTTCTTCAGTGGCCCTGGGCACTCTGTAGGGA 739  
600 AGACACTAAGAAAGAAACGACTTATTTCTTCAGTGGCCCTGGGCACTCTGTAGGGA 659

620 CCGGCAAGACAAACCGCCCTTAAGACGTTCTTGGAATTTGGGAMCAATTTGACCCCTC 679  
540 CCGGCAAGACAAACCGCCCTTAAGACGTTCTTGGAATTTGGGAMCAATTTGACCCCTC 599

560 GGACCCCGGCTTACATTTTGGACCAACGAGCATCCMAAGTGATGGGAAACGTTTC 619  
480 GGACCCCGGCTTACATTTTGGACCAACGAGCATCCMAAGTGATGGGAAACGTTTC 539

440 TTGGAATCCGTGAGGCAACGAACTCCAGGTGAGAGATACGAGCTTGCACCATCTTG 499  
361 TTGGAATCCGTGAGGCAACGAACTCCAGGTGAGAGATACGAGCTTGCACCATCTTG 419

500 GAAGCGGCTGCTACCTCTTCGAGAGTGGTTCCACCACTCTTGGGAGCTCTGTGAGCA 559  
420 GAAGCGGCTGCTACCTCTTCGAGAGTGGTTCCACCACTCTTGGGAGCTCTGTGAGCA 479

380 TGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGGCCCAAGATTTCCATTTCC 439  
301 TGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGGCCCAAGATTTCCATTTCC 360

440 TTGGAATCCGTGAGGCAACGAACTCCAGGTGAGAGATACGAGCTTGCACCATCTTG 499  
361 TTGGAATCCGTGAGGCAACGAACTCCAGGTGAGAGATACGAGCTTGCACCATCTTG 419

500 GAAGCGGCTGCTACCTCTTCGAGAGTGGTTCCACCACTCTTGGGAGCTCTGTGAGCA 559  
420 GAAGCGGCTGCTACCTCTTCGAGAGTGGTTCCACCACTCTTGGGAGCTCTGTGAGCA 479

320 AGTGCCCTGGGTTTGTCTAATTTGAGCTGAACACTANTCACTGGGTTTCCATGGTTCTCTTC 379  
241 AGTGCCCTGGGTTTGTCTAATTTGAGCTGAACACTANTCACTGGGTTTCCATGGTTCTCTTC 300

380 TGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGGCCCAAGATTTCCATTTCC 439  
301 TGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACATGGGCCCAAGATTTCCATTTCC 360

440 TTGGAATCCGTGAGGCAACGAACTCCAGGTGAGAGATACGAGCTTGCACCATCTTG 499  
361 TTGGAATCCGTGAGGCAACGAACTCCAGGTGAGAGATACGAGCTTGCACCATCTTG 419

500 GAAGCGGCTGCTACCTCTTCGAGAGTGGTTCCACCACTCTTGGGAGCTCTGTGAGCA 559  
420 GAAGCGGCTGCTACCTCTTCGAGAGTGGTTCCACCACTCTTGGGAGCTCTGTGAGCA 479



ORGANISM	Homo sapiens	Db	1827	CACTCAATAAGGAAGAAGGTGAGAAAACAACCCAGCATTCACATATCCCAATAGGAAGCGAG	1886
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Qy	1355	GCAGTTCCTCAGTCTASACCTCATCTGGGGACACAGAAATCACTAACATGGAGATTTGGTG	1414
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	Db	1887	GCAGTGGCCCTTGAAACAACATAATTCGGACCCAG----TGGTACCTTTAGATGAGTGCA	1942
CONSRM	1 (bases 1 to 6400)	Qy	1415	CTGCAGACATTTGCTAACTTGTGTCTASAGGACTAAGGAAAACCTASGAGAAGAAATCTA	1474
TITLE	koehrer,k., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,	Db	1943	AAGAAACATTTTCAATATGTCTTATTAGAGGGCTTACAGAGGACCA-GAGCCAGCCCTC	2001
JOURNAL	Fobo,G., Han,M. and Wiemann,S.	Qy	1475	YGAATTACTCAATGATGTCCACCAATAACACAGGGGAAGGGAAGAAAATCTTACTGCTTTT	1534
COMMENT	The German cDNA Consortium	Db	2002	TTAATTACCCCAAACTCTCCATGATAGATCA-AAAACCTAGATGAAAATCCCTCAGCCCTT	2060
	Direct Submission	Qy	1535	CTGAGAGACATAAGGGAGGCATTTGAGGAAGCGTCTCTCTGTCACTGT-----	1583
	Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764	Db	2061	TTGGAAAGGCTGAGAGAGGCTGTAGTGAACAACACTTCTCTGTCTCTGTGAGGTATGTT	2120
	Neuherberg, GERMANY	Qy	1584	-----ACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC	1626
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	Db	2121	TATACCTCACATAAATTAATTCAGTTGAGGGACAGCTAATTTTGAAGCAAGTTTATTATC	2180
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;	Qy	1627	TCAGTTCAGCTGCAGACATTAGAA-AAAACTTTCAAAAGTCTCCGCTAGGCCCGGAGCAAAA	1685
	sequenced by BMFZ (Biomedical Research Center at the	Db	2181	TCAGACAGCCCTCGATATCAGAAGGAAGCTAGAGAAAACAGGCCCATGGGATTAGATAGCAC	2240
	Heinrich-Heine-University, Dueseldorff/Germany) within the cDNA	Qy	1686	CTTAGAAACCTTATGAACCTTGGCAACVTCGGTTTTTTTATATATAGATCAGGAGGAGCA	1745
	sequencing consortium of the German Genome Project. This clone	Db	2241	CTTACAAAACCTCTCCTGAAGGTAGCCACCTCGTCTTTTATATATAGGACAGGAGGAGGC	2300
	Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.	Qy	1746	GGCGGAAACAGACAAAACGGGATTAATAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGC	1805
	Please contact RZPD for ordering:	Db	2301	CCAGGGGAGGAGGAGGAAAACAAGAAAAGGC-----	2333
	http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686121167	Qy	1806	AAGTGGACTTTGGAGGCTCTGGAAGGAGGAAAAGCTGGGCAAAATTTGAATGCTAATAGGG	1865
	Further information about the clone and the sequencing project is	Db	2334	-AGAGACTCTGTTGGCCACTTTTGTATAAACCCAGTATATCCCAAGGTACACCTGCTAG--	2390
	available at http://mips.gsf.de/projects/cdna/.	Qy	1866	CTTGCTTCAGTGGGCTCTACAAGACACTTTTAAAAAAGATGTCCAAAGTAGAAGTAAGC	1925
FEATURES	Location/Qualifiers	Db	2391	-TTGTACTCAATATGGCAAACTAGGGCACTTTCAAAAGGAGGACTGCCAGCAGCAAGAAGA	2449
source	1. .6400	Qy	1926	CGCCCTTCGTCATGCCCTTTTTCAGGGAATCTCTGGAAGGCCCACTGCCCCAGGG	1985
gene	/organism="Homo sapiens"	Db	2450	AGCCACCTCAACCTGTCTCAGTATACATGTTGGGACCACTGGAAGGACACTGTCTCTGGGA	2509
CDS	/mol_type="mRNA"	Qy	1986	GACAAAGGCTTTTGAGTCAGAAGCCACTAACCCAGATGATCCAGCAGCAGCACTGAG---	2042
	/db_xref="taxon:9606"	Db	2510	AGCATAGTCTACTGAGTCCAGAGCCAGTCTCACAAATGGTCCAGCAGCACTGATGGGTCC	2569
	/clone="DKFZp686121167"	Qy	2043	-----GGTGGCTGGGGCAAGGCCATCCCATGCCATCACCTCTACAGAGCCCTGGG	2093
	/tissue_type="uterus endometrial, primary cell culture"	Db	2570	TGGGGCTCAACTCCCTGGCTCCAGTGGCTCAGACTGTCTTACTACCCGGGAGCAACGAG	2629
	/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host	Qy	2094	TATGCTTGACCATTTGAGGSC--CAGGAAGGTGTCTCTCGACACTGTGCGGTCTTCT	2150
	DH10B; sites SfiIA + SfiIB"	Db	2630	TGATTCTGGAGTTGAGGGAGGAGGTAGACTTCTTTCTGACACTGGAGCCAGCT-T	2688
	/dev_stage="adult"	Qy	2151	TAGTCTTACTCTTGTCTCCGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGT	2210
	/note="hypothetical protein"	Db	2689	CTGTTTCTCTCCAAATTCAGGCCCTTCGTTCTCTCC--CATAGCAACAACCATGATGGGCA	2745
	1. .6400	Qy	2211	CCNTAAGAGGGGAGTCACTAGATACTTTTCCAGGCCACTAAGTTATGAATCTGGGAGC	2270
	/gene="DKFZp686121167"	Db	2746	TCTCAGGAAGAGCTTTTAACTTTGATATTTTCTCAACCTATTAGTTGCACTCAGAGGAGCC	2805
	209. .1573	Qy	2271	TTTATCTTTTTCACATGCTTTTCTAAATATGCTTGAAGGCCCACTACTCTGTTAGGAG	2330
	/gene="DKFZp686121167"	Db	2806	--TACTGTTTACAGGGCTTTTATGTCATGCTGAAAGTCCCACTCTCTTTATTATAGATAG	2863
	/codon_start=1	Qy	2331	AGACATCTTACAAAAGCAGGGGCAATTTATACCTTGAAACATAGGAAGGAACACCCGT	2390
	/product="hypothetical protein"	Db			
	/protein_id="CAH18335.1"				
	/db_xref="GI:51476728"				
	/translation="MTWAAAATCAALLLWALTTGDLKLVEMMAGGTQITPLNDNV				
	TIFCNIFSOPLNITSMGITWFKVTFDKEVTFEFGDHOEAPFCAIYSPWRLKS				
	GDASRLQLOLEAGEVRCVAVTFLKAQCTVOLEVVASPASRLLLDOVGMKNEKD				
	YNCESGYPERAINITWETQKFPHPHIEISDVITGTIKMDGTFTNWTCLKLNS				
	QBDPTVQCVVRHASLHTPLKSNFTLTAARHSLSSETKDNFSLHWWPISFGVGLV				
	LLIVLPWKIKCNKSSAYTPLKILKHWNSFDITLKEHLIFPCTRAWPSYQLQDG				
	EAWPESGVNINTIOQLDVFRCQEGKWSEVPYQAFALRNPDLCQCCRIDPALLTV				
	TSKSIDDNSTKSEKQTPREHSDAVDPILPVPSPWPPPTTTTFLVSSQRPPTLL				
	LPLQ"				
ORIGIN	Query Match				
	Best Local Similarity 54.9%; Pred. No. 8.3e-237;				
	Matches 2156; Conservative 112; Mismatches 1531; Indels 130; Gaps 33;				
Qy	1057 GGCACGACAGAGTGCATGCGCTTTTCTCCAGACTTAAGCAATAAATAACAGAC		1116		
Db	1589 GGTGCCACTAGGTGTCAGATTCCTTTTATACAGGACCTTTGGGCAATAAGAGGGGAC		1648		
Qy	1117 TTAGGTAAATCTCAGATAAATCTGATGCTATATTGTTGTTTCAAGGTTTAGACAA		1176		
Db	1649 CTGGGCAAGTCTCTGATGA-CTGATAGATATATAGAGGCTTTTCAAAACCTAATCTAG		1707		
Qy	1177 TTTCTTTGATCTGACATGGAGAGATATATGCTGCTTAAATCAGACACTAACCCCAA		1236		
Db	1708 GTGTTTGACCTTACTTGGGGTGATGT-TATGTTGCTCTTAAACTCTTAACTACTACAGAG		1766		
Qy	1237 TGAGAGAAGTGCCACCA--TAATCGAGCTCAGAGTTTGGGATCTCTGGTATCTCAGT		1294		
Db	1767 AAACAGGTAGCCCTGAGGACAGCAAAAAATTACAGAGATGAACAACATGCTCTCTATAGC		1826		
Qy	1295 CAGGTCAATGGATGATGACAAACAGAGAAAGAAANATGATTCCTCCACAGGCCAGCAR		1354		

Db	2864	AGATATTCCAGCAGCTATGGGGGCCAACAATCTTTATGGCTCCAGGACAAATCTCGCC	---	2929
Qy	2391	TTGTTGNTCCCTGCTTGAGGAAGAAATTAATCTCTGAAGTCTGGCAACAGAGCAAT	2450	
Db	2921	-----TCCCTTAGTGGAAACCAAGTATTAAACCAGAGGTTTGGCAACTCAGGAAGAA	2974	
Qy	2451	ATGGACGAGCCAAAGAATGCCCCGTCCTGTTCCAGTTTAAACTAAAGGATTCACATCTCTTT	2510	
Db	2975	GTGGCCG-GACTATTAACCACTATACCAAGTCTGGATCCACCTCAAGGATCCCACTTTT	3033	
Qy	2511	CCCTACCAAGGCTAGTACCCCTCAGACCCCAAGGCCCAACAAGGATTCCTAAAGATTTGT	2570	
Db	3034	CCTAACAGAGACAATACCCCTGAGCGCAGAAAGTAGAAAGGGCTAGACGCCAGCAT	3093	
Qy	2571	AAGGACTTAAAGCCCAAGGCTTAGTAAAAACCATGATTAATCTCCCTGCAGTAATTCGGTA	2630	
Db	3094	GATAACCTAAGGCACAAGGCTCCTTTAGGCTTTGCAGTAGCCCTTGAACAACCCCAATA	3153	
Qy	2631	GTGGAATTGAGGAGCCACAGAAACCCAGTGGACAGTGGAGGGTTAGTCCAAAGATCTCAGGA	2690	
Db	3154	TTAGG-----GGTGCAAAATCCAATGGGAAGGGA-GACTAGTTTCAGGACCTCTGCC	3205	
Qy	2691	TTATCAATGAGGCGGTGTCTTTTATACCCAGCTGTAACCTAGCCCTATACGTGMYT	2750	
Db	3206	TTATTAT-GAGGTTGTGATTACGATTCATCCAGTAGATCCAGTCCCTATACTTGCCTA	3264	
Qy	2751	TCCCAATACAGAGGAAGCAGAGTGGTTTACASTCTCTGGACCTTTCAGGATGCTCTTCTC	2810	
Db	3265	ACTCAATATCCCAAGGGAATCAATGGTTTCACAGTTCTGGGCTTAAGGATGCTTTTTC	3324	
Qy	2811	TGCATCCTGTACATCTCTGACTCTCAATCTTGTGTGCTTTGAAGATATTCAAACCCA	2870	
Db	3325	TGTATACCTTTATACCTCGACTCCCAATACTTGTTTACCTTGAAGACCCCTCCAATCAA	3384	
Qy	2871	RCATCTCAACTCAGCTGGACTRTTTTACCCCAAGGGTTTCAGGATAGYCCCCATCTATTT	2930	
Db	3385	ACACCCAGTTTAACCTGGACTGTACTGCTCAGGGATTTAGAGATGCCCCCACTATTT	3444	
Qy	2931	GGCCAGGACTTAGCCCAAGACTTTGAGYCAATYMFCTATACCTGGACACTCTGTGCTTCRG	2990	
Db	3445	GGGCAAGTATTGTCAGAGATCTCTTCGAATCTTCCAATCTCAGGTTAAGTTTACAA	3504	
Qy	2991	TAGTGTAGATTATCTTTTTRGVGVCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAA	3050	
Db	3505	TATGTAGATGACATCTCTCTGTGCCCCCACTGAGGAAGTTTCTTGAGACGCACCTGAG	3564	
Qy	3051	GCRTCTTMAATTCCTCGCYACCTGTGGCTCAWGGTTTCAAAASABARGCTCARTCT	3110	
Db	3565	GCCTCTCTTAATTTCTTAGCTGGAAGGGAATAAAGGCTTCAAAATCTAAGGCTCAGCT	3624	
Qy	3111	TGCTCAGCAGGTTAAATACTTAGGCTTAARATTTATCCAAAGGCACCAAG-GCCCTCAG	3169	
Db	3625	TATCAGACTTTGGTAAAGTACTTAGTCTGGTCTTGTTCAGAGGGAACAAGACACTAGG	3684	
Qy	3170	TGAGGAAYRYATCCAGCTTATATCGGCTTATCTCTCATCYCAAAACCCCTAAGCAACTAG	3229	
Db	3685	CGAGGAGGATTAAATCCCAATTTCTCCCTTCTCTCCCCAGAACCCCTCAAGCAGCTGAG	3744	
Qy	3230	RGRRTCTTGGCBRTAAYAGYTTCTGCCGAATATGATTTCCCAAGTWTGGCRAAATAG	3289	
Db	3745	-GGATGCTTGGGCATTACAGSTTTTGTAGGTTTATAGA-TACCCAGGATATGATAAAATAG	3802	
Qy	3290	CCAGGYCATTTAWATACASTAATTAAGGAAACTCAGAAAGCCCAATACCCATTTTARTAGAT	3349	
Db	3803	GTTCGCCCTTTATTCACCTTATTAAGGAATTCAGTGGCTTAAACTCACTTCTCAGCT	3862	
Qy	3350	GGAYMCTGAAGYMRAGTGGCTTTCAGGCCCCCTTAAAGAGGCGCTTAAACCCCAAGYCCC	3409	
Db	3863	GGGAACTGGAAGGTCAAAAGGCCATTAAATCAGTTTAAACAAG--CCTGTCTTAAAGCACC	3920	
Qy	3410	AGTGTTHAAGYTTGCCAACRGGGCAAGACTTTTGTATATYRTCAAGAAACCAAGAAACAGRAA	3469	
Db	3921	AGCCCCAGCTCTTTCTGTAGGGAAGTCTTTAATCTTTTATGATCATATAAGGAA-AGGAA	3979	

QY	3470	YAGCTCTGGAGTCCTTACACAGRTCCRAGGAYGAGCTTGCAACCTYRTGGCRYACCTGA	3529
Db	3980	TGACCTCGGAGTTTAAATCAAGTTCAAGGCCAGCTCAACAGCCAGTAGATTACTTAA	4039
QY	3530	STAAGGAAATGATGTAGTGCAAGGGTTCRCYTCATTGTTTAYGGTAGTGGTGCGAG	3589
Db	4040	GTAGAGAGCTTGACTTTGGAGGCTTAAGATGCGCAGCTTGCCCTCTGAGCAGTTACAGTAG	4099
QY	3590	TAGCAGTYKTAGTATCTGAACGAGTTAAAATAATACAGGGRAGAGATCTTACTGTGTGGA	3649
Db	4100	TGGCCCTGCTAGTCCAGAGGCCAACAGTTTAAACCTTGGGAATAACCTGTTTATGCTTC	4159
QY	3650	CATCTCATGAGTGAAVYRGATYACTACTGCTAAAGGAGACTTTGTGGCTGTGAGCAACY	3709
Db	4160	ACATAAAATAATGTGGCAGGCTTGCTGTCTCTAAAGCGAGTCTCTGGCTAACAGACAGCT	4219
QY	3710	GTTTACTTAAATATCAGGCTCTATTACTTGAARGCCAGTGCTGCRACCTGTGCACTTTGTG	3769
Db	4220	GCCTCCTCAAGTATCAGACTTTGCTGTAGGAATCTGCAGTCCAGTTAAAACCTGCAC	4279
QY	3770	CAACTCTTAAACCCAGYCNCAATTTCTTCCAGACAATGAAGAAAGATAAAYATAACTGTG	3829
Db	4280	CTTGCCCTGATCCAGACACTTTCTCTCCAGAGGAACTGGAGAACTGACACATAATGTG	4339
QY	3830	AACAARTAAATTTCTCAAACTATGCCACTCGAGGGACCTTGTAGAGTTTCCVTTGACTG	3889
Db	4340	AAC-AGACTGTAGTACAACTTATGCCGTAGGAGGAGTCTCAAAGMAATTTCCCTAGAAA	4398
QY	3890	ATCCYGACCTTCAACTTGTATACTGATGGAAGTTCCTTTGTAGAAAAGAGACTTCGAAA	3949
Db	4399	ATCCAG-GCTGGACCCACTTTATGGATGGGAGCTCTTTTGTAGACAGGGAGTCTGAAA	4457
QY	3950	YVGGGTATGCACTGGTTCAGTGATAATGGAAATVTTGAAAAGTAATCCCTCACTCCAGGA	4009
Db	4458	GTGGGATATGCACTAGTACCCCTAATAACA---TTGAAAATGTGCTCTCTCACAGGC	4514
QY	4010	ACTAGTCTYAGTRGCAGAACTAATAGCCVTCAYTKGGGCACCTAGAAATTAGGAGAAGRA	4069
Db	4515	ACACACTCAATCAGTTTGAAGTATCCCTCTCAAGAGCAGCTTGAACCTAAGCAAGGA	4574
QY	4070	AAAAGGYAATATATATACAGACTCTPRATATGCTTACCTAGTCNTCCATGCCCATGMR	4129
Db	4575	AAGCAGCTAACATTTATCTGACTCCAAGTGTACCTTTCTAGTTCTCTCTTGCCCATGCT	4634
QY	4130	GCAATATGSRAGAAAGGGAATTCCTAACTTCYVAGRGAACACCTATCAMACATCAGAA	4189
Db	4635	GCCATCTGAAGGAAGGCACTTCCCTCACTGCTAATGGGTCTTGCAATTAATACCATCAA	4694
QY	4190	GCCATTAGGARATTTATTAYTGGWGTTACAGAAACCTARAGAGGTGGMAGTCTTACACTGC	4249
Db	4695	GACATCAACAGGCTATTATCCTCAGTGTTCCTCCCATGAGAAATTAGTGGTTCAGCATTCG	4754
QY	4250	YGGGTTCATCAAAAAGGAAAGRAAGGGAATAASAGRGAAYTGCCAAAGCAKATATTGAA	4309
Db	4755	AGGGGATCAGAAAGCGGTAGATGAATAGCCAAAGGAAACAGGCTAGCAGACCAGAA	4814
QY	4310	GCMAAAGAGCTGCAAGGCAAGGACCTCTCATTTAGAAATGCTTTATTAACCTTCCCTTAGTA	4369
Db	4815	GCTAAGTCAGCAGTAAAGAGGCCCAAGTCCCAAAACACTTTGAGGCCCTCTCTGATTTGG	4874
QY	4370	TAGGTTAATCCCTTCCGGGAAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATCGGGA	4429
Db	4875	GAGGGCTA---CATAGGGAATAAGCCTCAGTATTCCTTACAGAGATAGAAATGGGCC	4931
QY	4430	ACCTCAGAGG--CAGTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGGGAAATATC	4487
Db	4932	ACCTCTCGAGGTATACCTTTCAACCCCTCAGATGGTTACAATCAGAAGATTGCAAAAGTAC	4991
QY	4488	TTTTGCCCTGCMACTATCCAATGGAAATTACTTAAAACCTTTGATCAAAACCTTTCACTTAG	4547
Db	4992	GCTTGCAGGCTCCAGCAATGGAAGATCTTAAATCTCTCACTGAGCCTTTCACTTAG	5051



clone CS0D1051YM13 5-PRIME, mRNA sequence.

ACCESSION  
BX389656  
VERSION  
BX389656.1  
KEYWORDS  
EST.  
SOURCE  
GI:30462930  
ORGANISM  
Homo sapiens (human)

Homosapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1. (bases 1 to 931)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0BAP0272B05\_AP02531\_1&c=4215.r

#### FEATURES

Location/Qualifiers

1..931  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1051YM13"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

#### ORIGIN

Query Match 11.0%; Score 820.8; DB 4; Length 931;  
Best Local Similarity 97.1%; Pred. No. 5.5e-229;  
Matches 899; Conservative 8; Mismatches 12; Indels 7; Gaps 7;  
QY 5848 AAAATGATTAATCCTAGTTGTCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTC 5907  
Db |||||  
2 AAAAGATTAACTCTAGTTGTCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTC 61  
QY 5908 ACCCAAACTGGTATGTCGTATGGGGGTGGAGTTCAGATCAGGAGAGAGAAAAACATGTA 5967  
Db |||||  
62 ACCCAAACTGGTATGTCGTATGGGGGTGGAGTTCAGATCAGGAGAGAGAAAAACATGTA 121  
QY 5968 AAGAAGTAGTATCCCAACTCACCGGGGTACATGGCACCTCTAGCCCTACAGGACTA 6027  
Db |||||  
122 AAGAAGTAGTATCCCAACTCACCGGGGTACATGGCACCTCTAGCCCTACAGGACTA 181  
QY 6028 GATCTCTCAAAACTACATGAAACCCCTCCGTACCATCTCGCTCGTGAAGCCCTATTTAAT 6087  
Db |||||  
182 GATCTCTCAAAACTACATGAAACCCCTCCGTACCATCTCGCTCGTGAAGCCCTATTTAAT 241  
QY 6088 ACCACCTCTCAGGGCTCCATGAGGTCCTGGGCCAAAACCTACTAACTGTTGGATATGC 6147  
Db |||||  
242 ACCACCTCTCAGGGCTCCATGAGGTCCTGGGCCAAAACCTACTAACTGTTGGATATGC 301  
QY 6148 CTCCTCCCTGACTTCGAGCCATATGTTCAATCCCTGTACCTGAAACAATGGAACTTC 6207  
Db |||||  
302 CTCCTCCCTGACTTCGAGCCATATGTTCAATCCCTGTACCTGAAACAATGGAACTTC 361  
QY 6208 AGCAGAGAATAAACACCACTTCCGTTTGTAGTAGGACCTCTGTTTCCATSTGGAAATA 6267  
Db |||||  
362 AGCAGAGAATAAACACCACTTCCGTTTGTAGTAGGACCTCTGTTTCCATSTGGAAATA 421  
QY 6268 ACCCATACCTCAAACTCCTGTTGTAAATTTAGCAATACTACATACCAACCACTCC 6327

Db |||||  
422 ACCCATACCTCAAACTCCTGTTGTAAATTTAGCAATACTACATACCAACCACTCC 481  
QY |||||  
6328 CAATGATCAGGTGGGTAACTCTCCACACAATAAGTCTGCTACCTCAGGAATATTT 6387  
Db |||||  
482 CAATGATCAGGTGGGTAACTCTCCACACAATAAGTCTGCTACCTCAGGAATATTT 541  
QY |||||  
6388 TTTGCTGTGGTACCTCAGCTATCGTGTGTTGAATGGCTCTTTCAGAACTATGTGCTTC 6447  
Db |||||  
542 TTTGCTGTGGTACCTCAGCTATCGTGTGTTGAATGGCTCTTTCAGAACTATGTGCTTC 601  
QY |||||  
6448 CTCTCATTTCTTAGTGCCCTATGACCATCTACACTGAAACAGATTATATACAGTTATGTC 6507  
Db |||||  
602 CTCTCATTTCTTAGTGCCCTATGACCATCTACACTGAAACAGATTATATACAGTTATGTC 661  
QY |||||  
6508 ATATCTAAGCCCGCAACAAAAGAGTACCCATTCTTCTTTTGTATAGGAGCAGGAGTG 6567  
Db |||||  
662 ATATCTAAGCCCGCAACAAAAGAGTACCCATTCTTCTTTTGTATAGGAGCAGGAGTG 721  
QY |||||  
6568 CTAGGTGCATAGGTACTGGCATTTGGCGGTATCACAACTCTACTCAGTTTCTACTACAAA 6627  
Db |||||  
722 CTAGGTGCATAGGTACTGGCATTTGGCGG-ATCACAACTCTACTCAGTTTCTACTACANA 780  
QY |||||  
6628 CTATCTCAAGAACTAAATGGGAC-ATGGAAACGGGTGCGGACT-CCCTGGTCACCTTGC 6685  
Db |||||  
781 CTATCTCAAGAACTAAATGGGACAAATGTAACGGGTGCGGACTCCCTGGTCACCTTGC 840  
QY |||||  
6686 AAGATCAACTTAAC-TCCCTTAGCAGCAGTA-GTCCTTCRAAATCGAAG-AGCTTTAGACT 6742  
Db |||||  
841 CAGATCAACTTACCTTCCCTTAGCAGCAGGAGGTCTTCAAAATCGAAGAGCTTTAGCCT 900  
QY |||||  
6743 YGCTAA-CGCTGAGAGAGGGGGAAC 6767  
Db |||||  
901 TGTACCCGCTGCAAGAGGGGGAAC 926

Search completed: August 2, 2006, 01:42:15  
Job time : 30592 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:12:25 ; Search time 1319 Seconds  
(without alignments)  
10755.684 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatcggtataataacc.....tattaatacttgactgtgr 7582

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCOMB COMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP COMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6448	86.1	13537	3	US-09-949-016-17382 Sequence 17382, A
2	6436	85.9	10499	3	US-09-719-554-3 Sequence 3, Appli
3	6124.6	81.8	8523	3	US-09-573-080A-21 Sequence 21, Appl
4	6124.6	81.8	8523	5	US-09-854-867-21 Sequence 21, Appl
5	2994.2	40.0	77772	3	US-09-949-016-17417 Sequence 17417, A
6	2994.2	40.0	77997	3	US-09-949-016-12249 Sequence 12249, A
7	2888.6	38.6	2946	3	US-09-175-928-3 Sequence 3, Appli
8	2615.4	34.9	168394	3	US-09-949-016-13002 Sequence 13002, A
9	2550.4	34.1	2599	3	US-09-719-554-1 Sequence 1, Appli
10	2282.4	30.5	2784	3	US-09-719-554-4 Sequence 4, Appli
11	2220.4	29.6	145320	3	US-09-949-016-15858 Sequence 15858, A
12	2216	29.6	46340	3	US-09-719-554-66 Sequence 66, Appl
13	2204	29.4	2763	3	US-09-949-016-5640 Sequence 5640, Ap
14	2168	28.9	8399	3	US-09-573-080A-26 Sequence 26, Appl
15	2168	28.9	8399	5	US-09-854-867-26 Sequence 26, Appl
16	2032	27.1	2055	3	US-09-719-554-22 Sequence 22, Appl
17	1837.4	24.5	2304	3	US-08-979-847B-87 Sequence 87, Appl
18	1507.4	20.1	276687	3	US-09-949-016-13840 Sequence 13840, A
19	1464.6	19.6	1799	3	US-09-719-554-5 Sequence 5, Appli
20	1452.4	19.4	2364	3	US-08-979-847B-88 Sequence 88, Appl
21	1415.8	18.9	2391	3	US-08-691-563C-57 Sequence 57, Appl
22	1415.8	18.9	2391	3	US-08-374-766-57 Sequence 57, Appl
23	1415.8	18.9	2391	3	US-08-979-847B-53 Sequence 53, Appl

24	1341.8	17.9	8919	3	US-09-573-080A-54 Sequence 54, Appl
25	1341.8	17.9	8919	5	US-09-854-867-54 Sequence 54, Appl
26	1324.2	17.7	1481	3	US-08-979-847B-105 Sequence 105, App
27	1312.4	17.5	1600	3	US-08-979-847B-207 Sequence 207, App
28	1301.8	17.4	1600	3	US-08-979-847B-206 Sequence 206, App
29	1282.4	17.1	1597	3	US-08-979-847B-205 Sequence 205, App
30	1200.2	16.0	3910	3	US-09-120-653D-1 Sequence 1, Appli
31	1195.4	16.0	1326	3	US-09-719-554-2 Sequence 2, Appli
32	1130.6	15.1	1740	3	US-09-719-554-61 Sequence 61, Appl
33	1087.2	14.5	1511	3	US-08-979-847B-117 Sequence 117, App
34	1081.4	14.4	1774	3	US-09-719-554-17 Sequence 17, Appl
35	1069.4	14.3	1577	3	US-08-691-563C-89 Sequence 89, Appl
36	1069.4	14.3	1577	3	US-09-374-766-89 Sequence 89, Appl
37	1069.4	14.3	1577	3	US-08-979-847B-83 Sequence 83, Appl
38	996.8	13.3	2330	3	US-09-120-653D-4 Sequence 4, Appli
39	954.2	12.7	1080	3	US-09-719-554-28 Sequence 28, Appl
40	950	12.7	1489	3	US-09-719-554-6 Sequence 6, Appli
41	947.8	12.7	1167	3	US-08-691-563C-61 Sequence 61, Appl
42	947.8	12.7	1167	3	US-09-374-766-61 Sequence 61, Appl
43	947.8	12.7	1167	3	US-08-979-847B-57 Sequence 57, Appl
44	932.4	12.4	1158	2	US-08-471-724-1 Sequence 1, Appli
45	932.4	12.4	1158	2	US-08-471-969-1 Sequence 1, Appli

## ALIGNMENTS

### RESULT 1

US-09-949-016-17382  
; Sequence 17382, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17382  
; LENGTH: 13537  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17382

Query Match 86.1%; Score 6448; DB 3; Length 13537;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGGAACGTTCCCGCAACACACAAACGCGCTTAAGACGCTATTCTCGARAATTGGGA 665

4610 GATGGGAACGTTCCCGCAACACACAAACGCGCTTAAGACGCTATTCTCGARAATTGGGA 4669

QY 666 MCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATCTCTGCAGTCCGCCCTG 725

4670 CCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATCTCTGCAGTCCGCCCTG 4729

QY 726 GCACCTCTGAGGAAGTATAAATTATAACCACTTACAGCTAGACATCTTTTGTAGAA 785

4730 GCACCTCTGAGGAAGTATAAATTATAACCACTTACAGCTAGACATCTTTTGTAGAA 4789

QY 786 AAGGCAAAATGAGTGAAGTGCATTAAGTACAACTTCTTTTTCATTAGAGCACTCAC 845

4790 AAGGCAAAATGAGTGAAGTGCATTAAGTACAACTTCTTTTTCATTAGAGCACTCAC 4849

QY 846 AATTATGAAAAAGTGTGATTATTATGCCCTACAGGAAGCCTTCAGAGTCTTACCTCCCTATC 905

Db	4850	AAATATGTAATAAAGTGTGATTTATGCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	5919	CGCCCCCTCTCCATGCCCCCTTATTTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG	5978
Qy	906	CCAGCATCCCGGACTCTTTCCCAATTAATAAGGACCCCTTTCAACCCAAATGTTCCAA	1985	GGACAAAGGTCTTTGAGTCAAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	2044
Db	4910	CCAGCATCCCGGACTCTTTCCCAACTAATAAGGACCCCTTTCAACCCAAATGTTCCAA	5979	GGACAAAGGTCTCTGAGTCAAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	6038
Qy	966	AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATATGA	2045	TGCTCGGGCAAGCGGCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTGACC	2104
Db	4970	AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATATGA	6039	TGCTCGGGCAAGCGGCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTGACC	6098
Qy	1026	CCCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCGACGACAGTGCGCTTCTTCTTTC	2105	ATTGAGGCGCAGGAAGGTTGTCTCTCGGACACTGGTGGGGTCTTCTTCTTCTTCTTCTTC	2164
Db	5030	CCCCCT-CCAAAGCAGTGGGAGGAGAAATTCGGCCGACGACAGTGCGCTTCTTCTTTC	6099	ATTGAGGCGCAGG-AGGTTGTCTCTCGGACACTGGTGGGGTCTTCTTCTTCTTCTTCTTC	6157
Qy	1086	TCCTCCAGACTTAAAGCAATAAAGACAGACTTAGGTAAATTTCTCAGATAAATCTGATGG	2165	TGTCCCGGACAACCTGTCTCTCAGATCTCTCACTATTTCTGAGGGGTCTCNNTAAGACGGGA	2224
Db	5089	TCTCCAGACTTAAAGCAATAAAGACAGACTTAGGTAAATTTCTCAGATAAATCTGATGG	6158	TGTCCCGGACAACCTGTCTCTCAGATCTCTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA	6215
Qy	1146	CTATATGRTGTTTACAAAGGTTTAGGACAATTTCTTTGATCTGACATGAGAGATATATA	2225	GTCACTAGATACTTTTCCAGGCCACTAAGTTATGAACCTGGGGAGCTTTATTTCTTTTTCAC	2284
Db	5149	CTATATGATGTTTACAAAGGTTTAGGACAATTTCTTTGATCTGACATGAGAGATATA-A	6216	GTCACTAGATAC-TTCTCCAGGCCACTAAGTTATG-ACTGGGGAGCTTTATTTCTTTTTCAC	6273
Qy	1206	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCACCACTAATCTGAGCCT	2285	ATGCTTTTCTTAATATGCTTGAAGCCCCCACTACCTTGTGTAGGGAGAGACATTTCTAGCAA	2344
Db	5208	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCACCACTAATCTGAGCCT	6274	ATGCTTTTCTTAATATGCTTGAAGCCCCCACTACCTTGTGTAGGGAGAGACATTTCTAGCAA	6333
Qy	1266	GAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACACAGAAGG	2345	AAGCAGGGGCGCATATACACCTGAACATAGGAGAAAGAACCCCTTTGTTGTGTCNCCCTG	2404
Db	5268	GAGGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAAACAG-AGG	6334	AAGCAGGGGCGCATATACACCTGAACATAGGAGAAAGAACCCCTTTGTTGTG-CCCCCTG	6392
Qy	1326	AAAGANAATGATTTCCCAAGCCGACGARGAGGTTCCCAAGTCTASACCTCATTTGGGAC	2405	CTTGAGGAAGGAAATTAATCTCTGAGTCTGGGCAACAGAGGACATATGAGCAGC-CAAA	2464
Db	5325	AAAGAGAATGATTTCCCAAGCCGACGARGAGTTCACCAAGTCTAGACCTCATTTGGG---	6393	CTTGAGGAAGGAAATTAATCTCTGAGTCTGGGCAACAGAGGACATATGAGCAG-CAAA	6451
Qy	1386	ACAGAAATCAGTAACATGGAGATTTGGTCTCAGACATTTTCTTAATTTGCTGTASAA	2465	GAATCCCCGCTCTCTGTTCAAGTTAAACTAAAGGATTCCTCTTCTTCCCTACCAAGGCA	2524
Db	5382	ACACAGAATCAGAACAAT-GGAGATTTGTTGCTCGAGACATTTGCTTAATTTGTTGTAGAA	6452	GAATCCCCGCTCTCTGTTCAAGTTAAACTAAAGGATTCCTCTTCTTCCCTACCAAGGCA	6511
Qy	1446	GGACTAAGGAAACTASGAAGAAATCTAYGAATTAATCAATGATGTCCACCATAAACACA	2525	GTACCCCCCTCAGACCCCAAGGCCCAACAAAGGATTCCAAAGATTTGTTAAGGACTTAAAGC	2584
Db	5441	GGACTAAGGAAACTTAGGAAG-AGTCTATGATTAATCAATGATGTCCACCATAAACACA	6512	GTACCCCCCTCAGACCCCAAGGCCCAACAAAGGACTTCCAAAAGATTTGTTAAGGACTTAAAGC	6571
Qy	1506	GGGGAAGGGAAGAAATCTCTGCTTTCTGGAGAGACTAAGGGAGGCAITTGAGGAAGC	2585	CCAGGCTTAGTAAACCATGCATAACTCCCTGCAAGTAACTCCGTAGTAACTCCGTAGTGAGGAG	2644
Db	5500	-GGGAAGGGAAGAAATCTCTGCTTTCTGGAGAGACTAAGGGAGGCAITTGAGGAAGC	6572	CCAGGCTTAGTAAACCATGCATAACTCCCTGCAAGTAACTCCGTAGTAACTCCGTAGTGAGGAG	2644
Qy	1566	GTGCTCTCTGTCACTGACTCTTCTGAAGGCCCACTAATCTTAAAGCGTAAAGTTATCA	2645	CACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGATCTCAGGATTAATCAATGAGGC	2704
Db	5559	GTGCTCTCTGTCACTGACTCTTCTGAAGGCCCACTAATCTTAAAGCGTAAAGTTATCA	6625	TACAGAAACCCACAGACAGTGA-GGTTAGTGCAGATCTCAGGATTAATCAAT-GAGGC	6682
Qy	1626	CTCAGTCAGCTGACACATTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA	2705	GTTGCTCTTTTATACCCAGCTGTACCTAGCCCTTATACCTGTGMYTTCCCAAATACCAGA	2764
Db	5619	CTCAGTCAGCTGACACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA	6683	TGTTGTTCTCTATAGCCAGCTGTACCTAGCCCTTATACCTGTGCTTTTCCCAAATACCAGA	6742
Qy	1685	ACTTAGAAACCTATTTGAATTTGGCAACTTCGGTTTTTTTATATAGAGATCAGGAGGAGC	2765	GGAAACAGAGTGTTTTACSTCCCTGACCTTMAAGGATGCTTCTTCTGATCCCTGTACA	2824
Db	5679	ACTTAGAAACCTATTTGAATTTGGCAACTTCGGTTTTTTTATATAGAGATCAGGAGGAGC	6743	GGAAACAGAGTGTTTTACAGTCTTCCGACCTTCAGAGTGCCTTCTTCTGATCCCTGTACA	6802
Qy	1745	AGGCGGAACAGGACAAACGGGATTAATAAAGGCGCCAGCTTTAGTCATGACCTCAGG	2825	TCCTGACTCTCAATTTCTTGTGTTGCTTGAAGATACCTTCAAAACCCACATCTCAACTCAC	2884
Db	5739	AGGCGGAACAGGACAAACGGGATTAATAAAGGCGCCAGCTTTAGTCATGACCTCAGG	6803	TCCTGACTCTCAATTTCTTGTGTTGCTTGAAGATACCTTCAAAACCCACATCTCAACTCAC	6862
Qy	1805	CAAGTGAATTTTGGAGCTCTGGAAAGGAAAGCTGGGCAAAATTTGAATGCTTAATAGG	2885	CTGGAATTTTAAACCCAGGGTTTCAAGGATAGYCCCAATCTATTTTGGCAGGCAATTAGC	2944
Db	5799	CAAGTGAATTTTGGAGCTCTGGAAAGGAAAGCTGGGCAAAATTTGAATGCTTAATAGG	6863	CTGGAATTTTAAACCCAGGGTTTCAAGGATAGTCCCAATCTATTTTGGCAGGCAATTAGC	6922
Qy	1865	GCTTGTCTCAGTGGCGTCTACAGGACAATTTTAAAGAGATTTGTCCTCAAGTGAAGTAAG	2945	CCAGACTTGAAGCTTCTTCTGACACTCTGACACTCTTGTCTTCTGATGATGATGATTT	3004
Db	5859	GCTTGTCTCAGTGGCGTCTACAGGACACTTTTAAAGAGATTTGTCCTCAAGTGAAGTAAG	6923	CCAGACTTGAAGCTTCTTCTGACACTCTGACACTCTTGTCTTCTGATGATGATGATTT	6980
Qy	1925	CGCCCCCTTCTGTCATGCCCCCTTATTTTCAAGGGAATCACTGGGAAGGCCCACTGCCCCAGG	3005	ACTTTTTRGCGCYRTTTCAGAAACCTTGTGCCATCAAGCCACCCCAAGCCTCTCTTAAATTT	3064
			6981	ACTTTTGGCGCCCATTCAGAAACCTTGTGCCATCAAGCCACCCCAAGCCTCTCTTCAATTT	7040





Db 9180 CTTAGCCAAATATCAACAAGTCTTAAAAATTAACAAGAACCTTATCCCTGAGAAGAGG 9239  
Qy 5284 GAAAGAACTATTCAACCCWGTGACATGATGATTAAGTCAAGTCCCTTCYCTCTAAATCC 5343  
Db 9240 GAAAGAACTATTTCACCCCTTGTGACATGATTAAGTCAAGTCCCTTCCTCTAAATCC 9299  
Qy 5344 CATCCCTAGATACATCTCTGGGAAGACCTACCCAGTCAATTTATATATACCCCACTGCGG 5403  
Db 9300 CATCCCTAGATACATCTCTGGGAAGACCTACCCAGTCAATTTATATATACCCCACTGCGG 9359  
Qy 5404 TTAAGAGTGGCTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCTCTGGATAC 5463  
Db 9360 TTAAGAGTGGCTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCTCTGGATAC 9419  
Qy 5464 CAAAGAACCTGAAATCAGAGACCAAGCTAGCTATTCCTGTGAACTCTAGAGGAT 5523  
Db 9420 CAAAGAACCTGAAATCAGAGACCAAGCTAGCTATTCCTGTGAACTCTAGAGGAT 9479  
Qy 5524 TCGCGCTGCTCTCAACAACCAAGAGGAAAGTAAATAAATCAATAATCCCCCATG 5583  
Db 9480 TCGCGCTGCTCTCAACAACCAAGAGGAAAGTAAATAAATCAATAATCCCCCATG 9538  
Qy 5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTVACCCTSTTTCACTCTCACTGCA 5643  
Db 9539 GSCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCTCTTCACTCTCACTGCA 9598  
Qy 5644 CCCCCTCCATGCGGTGATGACCAAGTACTCCCTTACCMAGAGTTCCTATGGAGATG 5703  
Db 9599 CCCCCTCCATGCGGTGATGACCAAGTACTCCCTTACCAAGAGTTCCTATGGAGATG 9658  
Qy 5704 CAGCTCCCGGAATATTGATGCCCTCGTATAGAGTCTTSTAAGGAGACCCCAACC 5763  
Db 9659 CAGCTCCCGGAATATTGATGCCCTCGTATAGAGTCTTSTAAGGAGAACCCCAACC 9718  
Qy 5764 TTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCAT 5823  
Db 9719 TTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCAT 9778  
Qy 5824 GCAATACTCATATTATGGACAGGAAATATGATTAATCTCTAGTGTCTCTGGAGGACTTGA 5883  
Db 9779 GCAATACTCATATTATGGACAGGAAATATGATTAATCTCTAGTGTCTCTGGAGGACTTGA 9838  
Qy 5884 GTCACTGTCTGTGGACTTACTTCAACCAATCTGATGCTCTGATGGGGGTGGAGTTCAA 5943  
Db 9839 GTCACTGTCTGTGGACTTACTTCAACCAATCTGATGCTCTGATGGGGGTGGAGTTCAA 9898  
Qy 5944 GATCAGGCAAGAGAAACATGTAAAGAAAGTAAATCTCCCAACTCAACGSGGTACATGCG 6003  
Db 9899 GATCAGGCAAGAGAAACATGTAAAGAAAGTAAATCTCCCAACTCAACGSGGTACATGCG 9958  
Qy 6004 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACCCAT 6063  
Db 9959 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACCCAT 10018  
Qy 6064 ACTCGCTGGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGTCTCGGCGCAA 6123  
Db 10019 ACTCGCTGGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGCTCTCGGCGCAA 10078  
Qy 6124 AACCTTACTAACTGTGTGGATATGCTCCCTGGAATCTTCARGCCATATGTTTCAATCCCT 6183  
Db 10079 AACCTTACTAACTGTGTGGATATGCTCCCTGGAATCTTCARGCCATATGTTTCAATCCCT 10138  
Qy 6184 GTACCTGAA CAATGGAACTTTCAGCACAGAAATAAACAACACTTCGTTTTAGTAGGA 6243  
Db 10139 GTACCTGAA CAATGGAACTTTCAGCACAGAAATAAACAACACTTCGTTTTAGTAGGA 10198  
Qy 6244 CCTCTGTTTCAATSTGAAATTAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
Db 10199 CCTCTGTTTCAATSTGAAATTAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 10258  
Qy 6304 AATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCCAACAAATA 6363

Db 10259 AATACTACATACAAACCAACTCCCAATGCATCAGGTGGTAACTCTCTCCACACAAATA 10318  
Qy 6364 GTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTTGAAT 6423  
Db 10319 GTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTTGAAT 10378  
Qy 6424 GGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTCTAGTGCCCCVATGRCCTATCTACACT 6483  
Db 10379 GGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTCTAGTGCCCCCTATGACCATCTACACT 10438  
Qy 6484 GAAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTC 6543  
Db 10439 GAAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTC 10498  
Qy 6544 CTTTTGTTATAGAGCAGGAGTCTAGGTGCACTAGGTACTTGGCATTTGGCGGTATCACA 6603  
Db 10499 CTTTTGTTATAGAGCAGGAGTCTAGGTGCACTAGGTACTTGGCATTTGGCGGTATCACA 10558  
Qy 6604 ACCTCTACTCAGTTCTACTACAAATCTCTCAAGAACTAAATGGGACATGGAACGGTCT 6663  
Db 10559 ACCTCTACTCAGTTCTACTACAAATCTCTCAAGAACTAAATGGGACATGGAACGGTCT 10618  
Qy 6664 GCCGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCAA 6723  
Db 10619 GCCGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCAA 10678  
Qy 6724 AATCGAAGAGTTTAGACTTACCGCTGABAGAGGGGAACTGTTTATTTTAGGG 6783  
Db 10679 AATCGAAGAGTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTTAGGG 10738  
Qy 6784 GAAAGATCTCTTATTTATGTTAACTCAATCCGGATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 10739 GAAGATGCTCTTATTTATGTTAACTCAATCCGGATCGTCACTGAGAAAGTTTAAAGAAAT 10798  
Qy 6844 CSAGATCGAATPACAACTGAKAGCAGAGCTTCGAAACACTTGGACCTCGGGCCCTCCTC 6903  
Db 10799 CGAGATCGAATPACAACTGAGCAGAGAGCTTCGAAACACTTGGACCTCGGGCCCTCCTC 10858  
Qy 6904 AGCBAATGGATGCGCTTGAATCTCCCTTCTTAGACCTCTAGCAGCTATTAATTTGCTA 6963  
Db 10859 AGCBAATGGATGCGCTTGAATCTCCCTTCTTAGACCTCTAGCAGCTATTAATTTGCTA 10918  
Qy 6964 CTCTCTTTGGACCTCTGATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATCGAA 7023  
Db 10919 CTCTCTTTGGACCTCTGATCTTTTAACTCTTGTAACTTTGTCTCTTCCAGAAATCGAA 10978  
Qy 7024 CTTGTRAACTCAAAATGGAGCCCAAGATGCAATGCAAGATCTACCGCAGACCC 7083  
Db 10979 CTTGTRAACTCAAAATGGAGCCCAAGATGCAATGCAAGATCTACCGCAGACCC 11038  
Qy 7084 CTGGACCGGCTGTTAGCCGACGATCTGATGTTAATGATCAATCAAGGACCCCTCTGAG 7143  
Db 11039 CTGGACCGGCTGTTAGCCGACGATCTGATGTTAATGATCAATCAAGGACCCCTCTGAG 11098  
Qy 7144 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAGCAGTTTAGAGCGT 7203  
Db 11099 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAGCAGTTTAGAGCGT 11158  
Qy 7204 SGTGGCCCAACTCCCAACAGCATTAGGTTTTCTCTGAGATGGGGGATCTGAGAGAC 7263  
Db 11159 SGTGGCCCAACTCCCAACAGCATTAGGTTTTCTCTGTTGAGATGGGGGATCTGAGAGAC 11218  
Qy 7264 AGGACTAGCTGATTTCTAGCTGATTAAGAACTCTTAAGCCTTAGSTGGAGGTGACC 7323  
Db 11219 AGGACTAGCTGATTTCTAGCTGATTAAGAACTCTTAAGCCTTAGSTGGAGGTGACC 11278  
Qy 7324 ACATCCACTTTAAACAGGGGCTTGGCAACTTGTATGATCACTGACCTCAATCAGAGGCTC 7383  
Db 11279 ACATCCACTTTAAACAGGGGCTTGGCAACTTGTATGATCACTGACCTCAATCAGAGGCTC 11338  
Qy 7384 ACTAAATGCTAAATTTAGGCAAGAGCAGGAGGTAAAGAAATAGCCCAATCATTTATTTGCTG 7443  
Db 11339 ACTAAATGCTAAATTTAGGCAAGAGCAGGAGGTAAAGAAATAGCCCAATCATTTATTTGCTG 11398

Qy	7444	AGACACAGCAGGAGGACCAATGATCGGGATATAAACCCCAAGTTCGACGCCGCAACGG	7503
Db	11399	AGACACAGCAGGAGGACCAATGATCGGGATATAAACCCCAAGTTCGACGCCGCAACGG	11458
Qy	7504	CAACCCCTTTGGGTCCCTCCCTTTCATGCGGAGCTCTGTTTTCATGCTATTTCACCTCT	7563
Db	11459	CAACCCCTTTGGGTCCCTCCCTTTCATGCGGAGCTCTGTTTTCATGCTATTTCACCTCT	11518
Qy	7564	ATTAAATCTTGCACCTGCR	7582
Db	11519	ATTAAATCTTGCACCTGCA	11537

RESULT 2

US-09-719-554-3

; Sequence 3, Application US/09719554

; Patent No. 6919438

; GENERAL INFORMATION:

; APPLICANT: ALLIEL, Patrick

; APPLICANT: PERIN, Jean-Pierre

; APPLICANT: RIEGER, Francois

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU

; FILE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS

; FILE REFERENCE: 200936USOPCT

; CURRENT APPLICATION NUMBER: US/09/719,554

; CURRENT FILING DATE: 2001-12-26

; PRIOR APPLICATION NUMBER: PCT/FR99/01513

; PRIOR FILING DATE: 1999-06-25

; NUMBER OF SEQ ID NOS: 122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 10499

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-719-554-3

Query Match	85.9%;	Score 6436;	DB 3;	Length 10499;
Best Local Similarity	95.7%;	Pred. No. 0;		
Matches 6680;	Conservative 195;	Mismatches 50;	Indels 54;	Gaps 28;
QY	606	GATGGGAAAGCTTCCCGCAAGACCAAAAAGCGCCCTTAAGACGTAATTTCTGGAAATTTGGGA	665	
DB	2953	GATGGGAAAGCTTCCCGCAAGACCAAAAAGCGCCCTTAAGACGTAATTTCTGGAAATTTGGGA	3012	
QY	666	MCAATTTGACCCCTCAGACACTAAGAAAAGAAAAGCACTTATTTCTTCTGCAGTGCGCCCTG	725	
DB	3013	MCAATTTGACCCCTCAGACACTAAGAAAAGAAAAGCACTTATTTCTTCTGCAGTGCGCCCTG	3072	
QY	726	GCACCTCTGAGGGAAGTATAAATTTATAACACCACTTTACAGCTAGACYCTCTTTGTAGAA	785	
DB	3073	GCACCTCTGAGGGAAGTATAAATTTATAACACCACTTTACAGCTAGACCTCTTTGTAGAA	3132	
QY	786	AAGGCAATGGAGTGAAGTGCCTAATAGTACAAACTTCTTTTCATTAAGAGACAACCTAC	845	
DB	3133	AAGGCAATGGAGTGAAGTGCCTAATAGTACAAACTTCTTTTCATTAAGAGACAACCTAC	3192	
QY	846	AAATTATGTAAAAAGTGTGATTTATATGCCCTACAGAGAGCCCTTCAGAGTCTACCTCCCTATC	905	
DB	3193	AAATTATGTAAAAAGTGTGATTTATGCCCTACAGAGAGCCCTTCAGAGTCTACCTCCCTATC	3252	
QY	906	CCAGCATCCCGGACTCCTTCCCCAMTAAATAAGGACCCGCCCTTCAACCCCAATTTGGTCCAA	965	
DB	3253	CCAGCATCCCGGACTCCTTCCCCAACTAATAAGGACCCGCCCTTCAACCCCAATTTGGTCCAA	3312	
QY	966	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGATGCCAATATTTCCCAATTTATGA	1025	
DB	3313	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGATGCCAATATTTCCCAATTTATGA	3372	
QY	1026	CCCTTCCCAAGCAGTGGGAGGAGAGAAATTCGGCCAGCCAGAGTGCATGTGCTTTTYY	1085	
DB	3373	CCCTTCCCAAGCAGTGGGAGGAGAGAAATTCGGCCAGCCAGAGTGCATGTGCTTTTTC	3431	

Qy	1086	TCTCCGAGACTTAAAGCAATAAACAAGACACTAGGTAAATTTCTCAGATAAATCTGTGATGG	1144
Db	3432	TCTCCGAGACTTAAAGCAATAAACAAGACACTTAGSTAAATTTCTCAGATAAATCTGTGATGG	3491
Qy	1146	CTATATTGRTGTTTTTAACAGGGTTAGGACAAATCTTTTGATCTGCATCGGAGAGATATATA	1205
Db	3492	CTATATTGATGTTTTTAACAGGGTTAGGACAAATCTTTTGATCTGCATCGGAGAGATATA-A	3550
Qy	1206	TGTCACCTGCTAAATCAGACACTAAACCCAAATGAGAGAAGTGCCAACATAACTGACGCT	1265
Db	3551	TGTCACCTGCTAAATCAGACACTAAACCCAAATGAGAGAAGTGCCACCATTAATCTGACGCT	3610
Qy	1266	GAGRGTTTTGCGGATCTCTGCTATCTCAGTCAGGCTCAATGGATANGGATACACAGAGG	1325
Db	3611	GAGRGTTTTGCGGATCTCTGCTATCTCAGTCAGGCTCAATG--ATAGATGACACAG-AGG	3667
Qy	1326	AAAGANAATGATTTCCCCACAGGCGCAGCAGCTTCCAGTCTTASACCTTCATTGGGGAC	1385
Db	3668	AAAGAGAATGATTTCCCCACAGGCGCAGCAGCTTCCAGTCTTAGACCTTCATTGGG--	3724
Qy	1386	ACAGAAATCAGTAAACATGGGAGATTTGGTGTGCGACATTTGCTTAATCTGTGTGCTASAA	1445
Db	3725	ACAGAAATCAGAAACAT-GGAGATTTGGTGTGCGACATTTGCTTAATCTGTGTGCTAGAA	3783
Qy	1446	GGCTAAGGAAACTASGAGAARTCTAAGANTTACTCAATGATGTCACCATTAACACACA	1505
Db	3784	GGCTAAGGAAACTATGGAAG-AAGTCTATGAATTAATCAATGATGTCACCATTAACACA	3842
Qy	1506	GGGAGGGGAAGAAATCTCTACTCCCTTTCTGAGAGACTTAAGGGAGGCAATTCAGGAAGC	1565
Db	3843	-GGGAGGGGAAGAAATCTCTACTCCCTTTCTGAGAGACTTAAGGGAGGCATTTGAGGAGC	3901
Qy	1566	GTGCTCTCTGTCACTGACTCTTCTTGAAGGCCAACTAATCTTTAAAGCGTAACTTTATCA	1625
Db	3902	GTGCTCTCTGTCACTGACTCTTCTTGAAGGCCAACTAATCTTTAAAGCGTAACTTTATCA	3961
Qy	1626	CTCAGTCAGCTGAGACATTTAG-AAAAAATCTCAAAGTCTGCCGTAGGCCCGGAGCAAA	1684
Db	3962	CTCAGTCAGCTGAGACATTTAGAAAAAATCTCAAAGTCTGCCGTAGGCCCGGAGCAAA	4021
Qy	1685	ACTTAGAAACCTTATTGAACTTGGCAACTTCGGTTTTTATATAGAGATCAGGAGGAGC	1744
Db	4022	ACTTAGAAACCTTATTGAACTTGGCAACTTCGGTTTTTATATAGAGATCAGGAGGAGC	4081
Qy	1745	AGCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG	1804
Db	4082	AGCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG	4141
Qy	1805	CAAGTGACTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAAAATGAAATGCTTAATAGG	1864
Db	4142	CAAGTGACTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAAAATGAAATGCTTAATAGG	4201
Qy	1865	GCTTGCTTCCAGTGGCGTCTACAAGACACTTTAAAAAGATTTGTCAGATAGAACTAAG	1924
Db	4202	GCTTGCTTCCAGTGGCGTCTACAAGACACTTTAAAAAGATTTGTCAGATAGAACTAAG	4261
Qy	1925	CCGCCCTTCTGTCATGCCCTTTATTTCGAAGGAATCACTGGAAGGCCCACTGCCCCAGG	1984
Db	4262	CCGCCCTTCTGTCATGCCCTTTATTTCGAAGGAATCACTGGAAGGCCCACTGCCCCAGG	4321
Qy	1985	GGAACAAAGTCTTTTGAGTCAAGACCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	2044
Db	4322	GGAACAAAGTCTTCTGAGTCAGAACCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	4381
Qy	2045	TGCTGGGGCAGCGCCATCCATGCCATCACTCCTCAGAGGCCCTGGGTATGCTTGACC	2104
Db	4382	TGCTGGGGCAGCGCCATCCATGCCATCACTCCTCAGAGGCCCTGGGTATGCTTGACC	4441
Qy	2105	ATTGAGGGCCAGGAAGTTGTCTCTCTGGACACTTGGTGGCGTCTTCTTAGCTTACTCTTC	2164
Db	4442	ATTGAGGGCCAGG-AGGTTGTCTCTCTGGACACTTGGTGGCGTCTTCTTAGCTTACTCTTC	4500
Qy	2165	TGTCGCCGCAACTGTCTCCAGATCTGTCACTATTCTCAGGGGGTTCNTAAGACGGGCA	2224

4501	Db	 TGTCCCGGCAAACTGTCTCTCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAAGACGGGCA	4558
2225	Qy	GTCACTAGATACATTTTCCAGCAGCACTAAGTTATGACATGGGAGCGCTTTATCTTTCTTCC	2284
4559	Db	 GTCACTAGATAC-TTCTCCCGCAGCCATAAGTTATG-ACTGGGAGCTTTATTTCTTTCTTCC	4616
2285	Qy	ATGCTTTTCTAAATTATCTGTTGAAAGCCCCACTACCTTGTGTAGGAGAGACATTTCTAGCAA	2344
4617	Db	 ATGCTTTTCTAAATTATGTTGAAAGCCCCACTACCTTGTGTAGGAGAGACATTTCTAGCAA	4676
2345	Qy	AAGCAGGGGCGCATATATACACTGTAACATAGGAGAGAAACCCCGTTTGTGTNCCCCTG	2404
4677	Db	AAGCAGGGGCGCATATATACACTGTAACATAGGAGAGAAACCCCGTTTGTGT-CCCTG	4735
2405	Qy	CTTGAGGAAGGATTAATCTCTGAAGTCTGGCAACAGAAAGACAATATGACAGCAAA	2464
4736	Db	CTTGAGGAAGGAAATTAATCTCTGAAGTCTGGCAACAGAAAGACAATATGACAG-CAAA	4794
2465	Qy	GAATGCCGCTCTGTTCAAGTTAACTAAAGGATTCACATCTCTTTCCCTTACCAAGGCA	2524
4795	Db	GAATGCCGCTCTGTTCAAGTTAACTAAAGGATTCACCTCTCTTTCCCTTACCAAGGCA	4854
2525	Qy	GTACCCCTCTCAGACCCAGGCCCAACAAAGATTCCAAAGATTGTTAAGGACTTAAAGC	2584
4855	Db	GTACCCCTCTCAGACCCAGGCCCAACAAAGACTCCAAAGATTGTTAAGGACTTAAAGC	4914
2585	Qy	CCAAGCTTTAGTAAACCATGTCATACTCCCTGCAGTAATTCGCTAGTGGATGAGAGG	2644
4915	Db	CCAAGGCTCTAGTAAAGCCATGCGATGTAACCCCTGCAGTACTCC-----AATTTTAGGAG	4967
2645	Qy	CACAGAAACCCAGTGGACAGTGGAGGGTGTAGTGCAGATCTCAGGATTTATCAATGAGGC	2704
4968	Db	TACAGAAACCCACAGACAGTGG-AGTTAGTGCAGATCTCAGGATTTATCAAT-GAGGC	5025
2705	Qy	CGTTGCTCTTTATPACCCAGCTGTACTAGCCCTTATATCTGTGMYTTCCAAATACCAGA	2764
5026	Db	TGTTGTTCTCTATAGCCAGCTGTACTAGCCCTTATACTCTGCTTTCCCAATACCAGA	5085
2765	Qy	GGAGCAGAGTGGTTTACSTCTCGGACCTTMMAGGATGCCTTCTCTGCACTCCCTGTACA	2824
5086	Db	GGAGCAGAGTGGTTTACSTCTCGGACCTTMMAGGATGCCTTCTCTGCACTCCCTGTACA	5145
2825	Qy	TCCTGACTCTCAATCTTTGTTGGCTTTGAAGATACTTCAAAACCAACATCTCAACTCAC	2884
5146	Db	TCCTGACTCTCAATCTTTGTTGGCTTTGAAGATACTTCAAAACCAACATCTCAACTCAC	5205
2885	Qy	CTGACTRTTTTACCCCAAGGGTTTCAGGATAGYCCCCCATCTATTTTGGCCAGGCATTTAGC	2944
5206	Db	CTGACTATTTTACCCCAAGGGTTTCAGGATAGTCCCCCATCTATTTTGGCCAGGCATTTAGC	5265
2945	Qy	CCAAGACTTGAGYCATYMTACATCTGGACACTCTTGTCTGTCGTAGTGTGATGATTT	3004
5266	Db	CCAAGACTTGAGCAATCTCTATACCTGGACA--CTTGTCTCTCGGTAGGTGATGATTT	5323
3005	Qy	ACTTTTGGCGCCVRYTTTCAGAAACCTTTGTGCATCAAGCCACCAAGCCTCTTTTAAATTT	3064
5324	Db	ACTTTTGGCGCCCATTCAGAAACCTTTGTGCATCAAGCCACCAAGCCTCTTTCAATTT	5383
3065	Qy	CCTCGYACCTGTGGCTACAWGGTTTCCAAACCAARAGCTCACTCTGTCTCAGCAGGTT	3124
5384	Db	CCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAGGCTCACTCTGTCTCAGCAGGTT	5443
3125	Qy	TAAATACTTAGGCTAARATTTATCCAAAGGACCAAGGCCCTGATGAGGAATRYATCCA	3184
5444	Db	----TACTTAGGGCTAAATTTATCCAAAGGACCAAGGCCCTCAGTGAAGAACATCA	5499
3185	Qy	GCCTTACTGGCTTATCTCATGYCAAAACCTTAAAGCAACTAAGRRRTTCTTGGCCT	3244
5500	Db	GCCTTACTGGCTTATCTCATGCCAAACCTTAAAGCAACTAAGGGATTTCTTGGCGT	5559
3245	Qy	AAYAGGYTTCTGCCGAAWATGGATTTCCCGAGTTGGCRAAATAGCCAGGYCATTAWATA	3304

QY	4385	CGGGAACACAGCCAGTACTACGACGAGGAGAAACAGAAATGGGAAACCTCACGAGG-CAG	4443
Db	6683	CGGGAACACAGCCAGTACTACGACGAGGAGAAACAGAAATGGGAAACCTCACGAGGACAG	6742
QY	4444	TTTTCTCCCTCGGAGCGGTAGCCACTGAAGAAGGAGAAATACTTTTGCTGCAACTAT	4503
Db	6743	TTTTCTCCCTCGGAGCGGTAGCCACTGAAGAAGGAGAAATACTTTTGCTGCAACTAT	6802
QY	4504	CCAAATGAATATCTTAAACCCCTTCATCAAACTTTCACTTAGGCAATGATAGCACCCA	4563
Db	6803	CCAAATGAATATCTTAAACCCCTTCATCAAACTTTCACTTAGGCAATGATAGCACCCA	6862
QY	4564	TCARATGGCCAAATCATTTTACTGACACAGCCCTTTCAAACACTATCAAGCABATAKT	4623
Db	6863	TCARATGGCCAAATCATTTTACTGACACAGCCCTTTCAAACACTATCAAGCABATAKT	6922
QY	4624	CAGGGCTGTGAATGTGCCARAAATAATCCCTGCCCTTATCGCCAAAGCTCCTTCAGG	4683
Db	6923	CAGGGCTGTGAATGTGCCARAAATAATCCCTGCCCTTATCGCCAAAGCTCCTTCAGG	6982
QY	4684	ARAAACAAACAGGCCATTTACCTGARAARACTGGCAACTGATTTTACCCCAAGCCCC	4743
Db	6983	ARAAACAAACAGGCCATTTACCTGARAARACTGGCAACTGATTTTACCCCAAGCCCC	7042
QY	4744	AAACCTCAGGGATTTCAATCTACTAGTCTGGGTARATATTTTACGGGTTGGGCARAG	4803
Db	7043	AAACCTCAGGGATTTCAATCTACTAGTCTGGGTARATATTTTACGGGTTGGGCARAG	7102
QY	4804	GCCTTCCCTGTAGGACAGAAAAAGGCCAAGAGGTAAATAAGGCACCTAGTTTCATGAANTA	4863
Db	7103	GCCTTCCCTGTAGGACAGAAAAAGGCCAAGAGGTAAATAAGGCACCTAGTTTCATGAANTA	7162
QY	4864	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	4923
Db	7163	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	7222
QY	4924	ACAGTAACCCAGGAGTATCCAGGGGTTAGGTATAGATATCACTTACCTGGGCTGA	4983
Db	7223	ACAGTAACCCAGGAGTATCCAGGGGTTAGGTATAGATATCACTTACCTGGGCTGA	7282
QY	4984	AGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA	5043
Db	7283	AGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA	7342
QY	5044	AAGCAACCCAGGAACCCACCTCACATGGCTGTCTGTGCTATAGCCTTAAAAAGA	5103
Db	7343	AAGCAACCCAGGAACCCACCTCACATGGCTGTCTGTGCTATAGCCTTAAAAAGA	7402
QY	5104	ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGGCC	5163
Db	7403	ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGGCC	7462
QY	5164	TTCAATACCAATGACCTGTGTGACCCAGACGCCAATCTAGTTGACAGATCACCT	5223
Db	7463	TTCAATACCAATGACCTGTGTGACCCAGACGCCAATCTAGTTGACAGATCACCT	7522
QY	5224	CCTTAGCCAAATATCAACAGTTCTTAAACATTACAGGAACCTATCCCTGAGAGAGG	5283
Db	7523	CCTTAGCCAAATATCAACAGTTCTTAAACATTACAGGAACCTATCCCTGAGAGAGG	7582
QY	5284	GAAGAACAATTATCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	5343
Db	7583	GAAGAACAATTATCCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	7642
QY	5344	CATCCCTAGATATCTCTGGAGAGCCCTACCCAGTCAATTTATTTATCCCACTGGCG	5403
Db	7643	CATCCCTAGATATCTCTGGAGAGCCCTACCCAGTCAATTTATTTATCCCACTGGCG	7702
QY	5404	TTAAAGTGGCTGGAGTGGAGTCTTGGATATCATCATCTTGAGTCAATCTCGATCTGC	5463
Db	7703	TTAAAGTGGCTGGAGTGGAGTCTTGGATATCATCATCTTGAGTCAATCTCGATCTGC	7762

QY	5464	CAAAAGGAACCTGAAATTCAGGAGACAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	5523
Db	7763	CAAAAGGAACCTGAAATTCAGGAGACAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	7822
QY	5524	TGCGCTGCTCTTCAAAACAACACAGGAGGAAGTAATAAATCATAAATCCCCCATG	5583
Db	7823	TGCGCTGCTCTTCAAAACAACACAGGAGGAAGTAATAAATCATAAATCCCCCATG	7881
QY	5584	GSCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTAAACCTCTTTTCACTCTCACTGCA	5643
Db	7882	GSCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTAAACCTCTTTTCACTCTCACTGCA	7941
QY	5644	CCCCCTCCATGCGCGTGTATGACAGTAGTCCCTCCCTYACMAGAGTTTCTATGGAGATG	5703
Db	7942	CCCCCTCCATGCGCGTGTATGACAGTAGTCCCTCCCTYACMAGAGTTTCTATGGAGATG	8001
QY	5704	CAGCGTCCCGGAAATATTCATGATGCCCATCGTATAGAGTCTTTTAAAGGGAAACCCCA	5763
Db	8002	CAGCGTCCCGGAAATATTCATGATGCCCATCGTATAGAGTCTTTTAAAGGGAAACCCCA	8061
QY	5764	TTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGSCCACTCTTTTGCATGCAT	5823
Db	8062	TTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGSCCACTCTTTTGCATGCAT	8121
QY	5824	GCMAATPACTATTTTGGACAGAAAAATGATTAATCTCTAGTTGCTCTGGAGGACTTGA	5883
Db	8122	GCMAATPACTATTTTGGACAGAAAAATGATTAATCTCTAGTTGCTCTGGAGGACTTGA	8181
QY	5884	GTCACTGTCTGTGGACTTACTTCAACCCAACTGCTATGCTGTATGATGGGGGTGGAGTTCAA	5943
Db	8182	GTCACTGTCTGTGGACTTACTTCAACCCAACTGCTATGCTGTATGATGGGGGTGGAGTTCAA	8241
QY	5944	GATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCAACTCACCSGGGTACATGGC	6003
Db	8242	GATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCAACTCACCSGGGTACATGGC	8301
QY	6004	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTTCGGTACCCAT	6063
Db	8302	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTTCGGTACCCAT	8361
QY	6064	ACTCGCTGGTAAAGCTATTTAATACCAACCTCACCTGGGCTCCATGAGGTCTCGGCCCAA	6123
Db	8362	ACTCGCTGGTAAAGCTATTTAATACCAACCTCACCTGGGCTCCATGAGGTCTCGGCCCAA	8421
QY	6124	AAACCTTACTAACTGTGTGGATATGCTCCCTGGAACCTTCARGCCATATGTTTCAATCCCT	6183
Db	8422	AAACCTTACTAACTGTGTGGATATGCTCCCTGGAACCTTCARGCCATATGTTTCAATCCCT	8481
QY	6184	GTACTGAAACAATGGAAACAACCTTCAGCAGAGAAATAAACACCACTTCCTGTTAGTAGGA	6243
Db	8482	GTACTGAAACAATGGAAACAACCTTCAGCAGAGAAATAAACACCACTTCCTGTTAGTAGGA	8541
QY	6244	CCTCTGTTTCAATSTGGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTTAGC	6303
Db	8542	CCTCTGTTTCAATSTGGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTTAGC	8601
QY	6304	AATACTACATACACAACCAACTCCCAATGCAATGCAATGCAATGCAATGCAATGCAATG	6363
Db	8602	AATACTACATACACAACCAACTCCCAATGCAATGCAATGCAATGCAATGCAATGCAATG	8661
QY	6364	GTCTGCTACCTCAGGAATAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6423
Db	8662	GTCTGCTACCTCAGGAATAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	8721
QY	6424	GGCTCTTACAGATCTATGTGCTTCTCTCAATTTTAGTGGCCCCCCTATGRCATCTACACT	6483
Db	8722	GGCTCTTACAGATCTATGTGCTTCTCTCAATTTTAGTGGCCCCCCTATGRCATCTACACT	8781
QY	6484	GAACAAGATTTATACAGTTATGTCTATCTTAAGCCCGCCCAACAAAGAGTACCCATCTT	6543
Db	8782	GAACAAGATTTATACAGTTATGTCTATCTTAAGCCCGCCCAACAAAGAGTACCCATCTT	8841
QY	6544	CCTTTTGTATAGGAGCAGGAGTGTAGTGCACTAGGTACTGGCAATTTGGCGGTATCACA	6603

```

Db      8842  CTTTTTGTATAGGACGAGGAGTCTAGTGTCACTAGGTACTGGCATTTGGCGGTATCACA 8901
QY      6604  ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 6663
Db      8902  ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 8961
QY      6664  GCGGACTCCCTGGTCACCTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCCTCBA 6723
Db      8962  GCGGACTCCCTGGTCACCTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCCTCAA 9021
QY      6724  AATCGAAGAGCTTTAGACTTGTCTTAACCGCTGAGAGAGGGGAACTCTTTTATTTTAGGG 6783
Db      9022  AATCGAAGAGCTTTAGACTTGTCTTAACCGCTGAGAGAGGGGAACTCTTTTATTTTAGGG 9081
QY      6784  GAAGAATGCTGTATTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843
Db      9082  GAAGAATGCTGTATTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 9141
QY      6844  CSAGATCGAATCAACGTTAKAGCAGAGAGCTTCGAAACACTGGACCTGGGGCCCTCCTC 6903
Db      9142  CGAGATCGAATCAACGTTAGCAGAGAGAGCTTCGAAACACTGGACCTGGGGCCCTCCTC 9201
QY      6904  AGCCRATGATGCCCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATATTGCTA 6963
Db      9202  AGCCAATGATGCCCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATATTGCTA 9261
QY      6964  CTCCTCTTTGGACCCCTGTATCTTTTACCTCTCTTTTAACTTTGTCTCTTCAGAAATCGAA 7023
Db      9262  CTCCTCTTTGGACCCCTGTATCTTTTAACTCTCTTTTAACTTTGTCTCTTCAGAAATCGAA 9321
QY      7024  GCTGTAAATCAAAATGAGGCCCAAGATGCGAGTCAAGACTAAAGATCTACCGGAGACC 7083
Db      9322  GCTGTAAATCAAAATGAGGCCCAAGATGCGAGTCAAGACTAAAGATCTACCGGAGACC 9381
QY      7084  CTGGACCGGCTGTAGCCCAAGATCTGATGTTTAAATGATCAATCAAGGACCCCTCTGAG 7143
Db      9382  CTGGACCGGCTGTAGCCCAAGATCTGATGTTTAAATGATCAATCAAGGACCCCTCTGAG 9441
QY      7144  GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAGCAAGTTAGACGGT 7203
Db      9442  GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAGCAAGTTAGACGGT 9501
QY      7204  SGTGGCCAACTCCGCCAAGCAGCTTAGGTTTCTGTTGATGAGTGGGGACTGAGAGAC 7263
Db      9502  C-TGGGCCAACTCCGCCAAGCAGCTTAGGTTTCTGTTGATGAGTGGGGACTGAGAGAC 9560
QY      7264  AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTTAGSTGGGAAGGTGACC 7323
Db      9561  AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTTAGSTGGGAAGGTGACC 9620
QY      7324  ACATCCACCTTTAAACAACGGGGCTTGCAACTTAGYTCAACCTGACCAATACAGAGGCTC 7383
Db      9621  ACATCCACCTTTAAACAACGGGGCTTGCAACTTAGCTTACACCTGACCAATACAGAGGCTC 9680
QY      7384  ACTAAATCTTAATTAGGCAAGACAGGAGGTTAAGAAATAGCCATCATYTTATTGCMTG 7443
Db      9681  ACTAAATCTTAATTAGGCAAGACAGGAGGTTAAGAAATAGCCATCATYTTATTGCMTG 9740
QY      7444  AGAGCACACGAGGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCCGCAACGG 7503
Db      9741  AGAGCACACGAGGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCCGCAACGG 9800
QY      7504  CAACCCCTTTGGGTCCCTCTCTTTGATGGAGCTCTGTTTTCATGCTATTTCACCTCT 7563
Db      9801  CAACCCCTTTGGGTCCCTCTCTTTGATGGAGCTCTGTTTTCATGCTATTTCACCTCT 9860
QY      7564  ATTAATCTTGTGACCTGCR 7582
Db      9861  ATTAATCTTGTGACCTGCA 9879

```

```

US-09-573-080A-21
; Sequence 21, Application US/09573080A
; Patent No. 682897
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: repeat region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: herv17
; PUBLICATION INFORMATION:
; PUBLICATION: Jurka, J; Malchewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSU: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE:
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-21

```

```

Query Match      81.8%; Score 6124.6; DB 3; Length 8523;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY      606  GATGGAAACGTTCCCGCAAGACAGAAACGCCCTTAAGACGTATTCTGGARAAATGGGA 665
Db      1923  GATGGAAACGTTCCCGCAAGACAGAAACGCCCTTAAGACGTATTCTGGARAAATGGGA 1982
QY      666  MCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCCGCTG 725
Db      1983  CCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCCGCTG 2042
QY      726  GCATCTCTGAGGAGATATAAATTATAACCATCTTACAGCTAGACTCTTTTGTAGAA 785
Db      2043  GCATCTCTGAGGAGATATAAATTATAACCATCTTACAGCTAGACTCTTTTGTAGAA 2102
QY      786  AAGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTTCATTAAGAGACAATCAC 845
Db      2103  AAGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTTCATTAAGAGACAATCAC 2162
QY      846  AATTATGTAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCCCTATC 905
Db      2163  AATTATGTAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCCCTATC 2222
QY      906  CCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCAATTTGTTCCAA 965
Db      2223  CCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCAATTTGTTCCAA 2282
QY      966  AAGGAGATAGACAAAGGGTAAACGTTGAAACAAAGAGTGCATATTTCCTCCCAATATGA 1025
Db      2283  AAGGAGATAGACAAAGGGTAAACGTTGAAACAAAGAGTGCATATTTCCTCCCAATATGA 2342
QY      1026  CCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCACGACGAGTGCATGTCYTTTYY 1085
Db      2343  CCCCT-CCAGCAGTGGGAGGAGAAATTCGGCCACGACGAGTGCATGTCCTTTTC 2401
QY      1086  TCTCCAGACTTAAGACAAATAAAGACAGACTTAGGTAAATTTCTCAGATAAATCTGATGG 1145
Db      2402  TCTCCAGACTTAAGACAAATAAAGACAGACTTAGGTAAATTTCTCAGATAAATCTGATGG 2461

```







Db 4589 CACTAATTAAGGAACTCAGAAAGCCATACCCATTTAGTAGATGGACAACCTGAAGTAG 4648  
QY 3365 AAGTGGCTTTCCAGGCCCTTAAGAAAGCCCTTAAACCCCAAGYCCAGAGTGTAAAGYTTGCC 3424  
Db 4649 AAGTGGCTTTCCAGGCC-----TAAACCAAGCCCAAGTGTAAAGTGTGCC 4694  
QY 3425 AACGGGCAAGACTTTTSTTVATAYRTACAGAAAAAACAAGAAACAGAAVAGCTCTGGAGTCC 3484  
Db 4695 AACGGGCAAGACTTTTTCATATGTACAG-AAAAACAAGAAATAGCTCTAGAGTCC 4753  
QY 3485 TTACACAGTCCRAGGGYAGCTTGCACACCVRTGGCRVACCTGASTAAGGAAAYTGATG 3544  
Db 4754 TTACACAGATCCGAGGGATGACTTGCACACTGTGGCATCTGACTAAGGAAATGTATG 4813  
QY 3545 TAGTGGCAAGGGTGTGRVYTCATTTGTAYGGGTAGTGGTGGCAGTAGCAGYKTAGTAT 3604  
Db 4814 TAGTGGCAAGGGTGTGACCTCATTTGTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT 4873  
QY 3605 CTGAAGCAGTTAAATATACAGGGBAGAGATCTTACTGTGTGGACATCTCATGATGTA 3664  
Db 4874 CTGAAGCAGTTAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTA 4933  
QY 3665 AYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTGCACAACTGTTTACTTAAATRTC 3724  
Db 4934 ATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTGCACAACTGTTTACTTAAATGTC 4993  
QY 3725 AGGCTCTATTACTTGAAGGCCAGTGTGCRACCTGTGCACTTGTGCACTTGTAAACCCAG 3784  
Db 4994 AGGCTCTATTACTTGAAGGCCAGTGTGCGACTGTGCACTTGTGCACTTGTAAACCCAG 5053  
QY 3785 YNCATTTCTTCAGACATGAGAAAGATARAAYATACTGTCNAAAPAAATTTCTC 3844  
Db 5054 CCACATTTCTTCAGACATGAGAAAGATARAACATAAATACTGTCAACAAGTAATTTCTC 5113  
QY 3845 AAACCTATGCCACTCGAGGGACCTTGTAGAGTTCCYTTGACTGATCCYGCACCTTCAAC 3904  
Db 5114 AAACCTATGCCACTCGAGGGACCTTGTAGAGTTCCYTTGACTGATCCCGACC-TCAAC 5172  
QY 3905 TTGTACTGATGGAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATCGAGTG 3964  
Db 5173 TTGTACTGATGGAGTTCTTTGTAGAAAAAGGACTTCGAAAAAGTGGGGTATCGAGTG 5232  
QY 3965 GTCAGTGATAATGGAATATTTGAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTR 4024  
Db 5233 GTCAGTGATAATGGAATATTTGAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTR 5292  
QY 4025 CGAGAACTAATAGCCVTCAYTKGGGCACTAGAAATTAGGAGAAAGAAAAAGGYYAAATATA 4084  
Db 5293 GCAGAACTAATAGCCCTCACTTGGGCACCTAGAAATTAGGAGAAAGAAAAAGGCAATATA 5352  
QY 4085 TATACAGACTCTAATATGCTYACCTAGTCCNTCCATGCCCATGMRCAATATGSAAGAA 4144  
Db 5353 TATACAGACTCTAATATGCTYACCTAGTCCNTCCATGCCCATGMRCAATATGSAAGAA 5412  
QY 4145 AGGGAATTCCTAATCTTCYAGRGAAACACCTATCCMAACATCAGGAAGCCATTAGBARATTA 4204  
Db 5413 AGGGAATTCCTAATCTTCYAGRGAAACACCTATCCMAACATCAGGAAGCCATTAGBARATTA 5472  
QY 4205 TTAYTGGCWGTACAGAAACCTARAGAGTGGMAGTCTTACACTGCYGGGGTCAATCANAAA 4264  
Db 5473 TTAYTGGCWGTACAGAAACCTARAGAGTGGMAGTCTTACACTGCYGGGGTCAATCANAAA 5532  
QY 4265 GGAAGGAAGGGAATASAAAGGAAATGCCCAAGCAKATATTTGAAGCMAAAAGAGCTGCA 4324  
Db 5533 GGAAGGAAGGGAATASAAAGGAAATGCCCAAGCAKATATTTGAAGCMAAAAGAGCTGCA 5592  
QY 4325 AGGCAGGACCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC 4384  
Db 5593 AGGCAGGACCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC 5652  
QY 4385 CGGGAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGG-CAG 4443

Db 5653 CGGGAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGGACAG 5712  
QY 4444 TTTTCTCCCTCGGAGCGGTAGCCACTGGAAGAGGGAATAATACCTTTTGCCTGCAACTAT 4503  
Db 5713 TTTTCTCCCTCGGAGCGGTAGCCACTGGAAGAGGGAATAATACCTTTTGCCTGCAACTAT 5772  
QY 4504 CCAATGGAAATTAATCTTAAACCCCTTCATCAAAACCTTTTCACTTAGGCATCGATAGCACCA 4563  
Db 5773 CCAATGGAAATTAATCTTAAACCCCTTCATCAAAACCTTTTCACTTAGGCATCGATAGCACCA 5832  
QY 4564 TCARATGGCCAAATCATTTATTTATCTGACACAGGCTTTTCAAAACTATCAAGCABATATK 4623  
Db 5833 TCAGATGGCCAAATCATTTATTTATCTGACACAGGCTTTTCAAAACTATCAAGCABATATK 5892  
QY 4624 CAGGCTGTGAATGTGCGARARAAATAATCCCTGCTGCTTATCGCCHAAGCTCTTTCAGG 4683  
Db 5893 CAGGCTGTGAAGTGTGCCAGAGAAATAATCCCTGCTGCTTATCGCCHAAGCTCTTTCAGG 5952  
QY 4684 ARAACAAARAACAGGCCATTAACCTGARAARAACCTGGCAACTGTATTTTACCCACAAGGCC 4743  
Db 5953 AGAACAAARAACAGGCCATTAACCTGARAARAACCTGGCAACTGTATTTTACCCACAAGGCC 6012  
QY 4744 AAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTTCAACGGGTGGCCARAG 4803  
Db 6013 AAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTTCAACGGGTGGCCARAG 6072  
QY 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGGTAATAAAGGCNACTAGTTTCATGAATA 4863  
Db 6073 GCCTTCCCTGTAGGACAGAAAGGCCCAAGGTAATAAAGGCNACTAGTTTCATGAATA 6132  
QY 4864 ATTCAGGATTTCCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTTCCAGGCC 4923  
Db 6133 ATTCAGGATTTCCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTTCCAGGCC 6192  
QY 4924 ACAGTAACCCAGGAGTATCCAGCGTTAGGTATACGATATCACTTACACTGCGCTGA 4983  
Db 6193 ACAGTAACCCAGGAGTATCCAGCGTTAGGTATACGATATCACTTACACTGCGCTGA 6252  
QY 4984 AGGCCACAGTCTCAGGGAAGGTGAGAAAAATGAATAAAYACTCAAAAGGACATCTAAA 5043  
Db 6253 AGGCCACAGTCTCAGGGAAGGTGAGAAAAATGAATAAAYACTCAAAAGGACATCTAAA 6312  
QY 5044 AAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTTGCCCTATAGCCCTTAAAAAGA 5103  
Db 6313 AAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTTGCCCTATAGCCCTTAAAAAGA 6372  
QY 5104 ATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 5163  
Db 6373 ATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 6432  
QY 5164 TTCAATACCAATGACCTTGTGTTGACCCCAAGACGCAACTTAGTTGGCAGACATCACCT 5223  
Db 6433 TTCAATACCAATGACCTTGTGTTGACCCCAAGACGCAACTTAGTTGGCAGACATCACCT 6492  
QY 5224 CCTTAGCCAAATATCAACAAGTTCTTAAAAATTAACAAGAACCTTCCCTCAGAGAGG 5283  
Db 6493 CCTTAGCCAAATATCAACAAGTTCTTAAAAATTAACAAGAACCTTCCCTCAGAGAGG 6552  
QY 5284 GAAAGAACTATTTCCACCCWMTGTCATGTTATAGTCAAGTCCCTTCYCTTAATTC 5343  
Db 6553 GAAAGAACTATTTCCACCCWMTGTCATGTTATAGTCAAGTCCCTTCYCTTAATTC 6612  
QY 5344 CATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCACTTATTTATTTACCCCACTGG 5403  
Db 6613 CATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCACTTATTTATTTACCCCACTGG 6672  
QY 5404 TTAAAGTGTCTGGAGTGGAGTCTTGGATACATCACTTTAGTCAAACTCCTGATACTGC 5463  
Db 6673 TTAAAGTGTCTGGAGTGGAGTCTTGGATACATCACTTTAGTCAAACTCCTGATACTGC 6732  
QY 5464 CAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACTCTTAGAGATT 5523  
Db 6733 CAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACTCTTAGAGATT 6792

Qy	5524	TGCGCTGCTCTTCAAACAACACGAGGAGAGTAACATAAATCATAAATCCCCATG	5583
Db	6793	TGCGCTGCTCTTCAAACAACACGAGGAGAGTAACATAAATCATAAATCCCATG	6851
Qy	5584	GSCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCCCTSTTTCACTCTCACTGCA	5643
Db	6852	GCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCCTCTTTCACCTCTCACTGCA	6911
Qy	5644	CCCCCTCATGSCCGCTGATGACCAAGTAGCTCCCTYACOMAGAGTTTCTATGGAGATG	5703
Db	6912	CCCCCTCATGSCCGCTGATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGGAGATG	6971
Qy	5704	CAGCGTCCCGAAATATTGATGCCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAACC	5763
Db	6972	CAGCGTCCCGAAATATTGATGCCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCAACC	7031
Qy	5764	TTCACTGCCACACCCATATATGCCCGGCAACTGCTATCACTGTGCCACTCTTTGCAATGCAT	5823
Db	7032	TTCACTGCCACACCCATATATGCCCGGCAACTGCTATCACTGTGCCACTCTTTGCAATGCAT	7091
Qy	5824	GCAAACTCATATTTCGACAGGAAATATGATTAATCCTTAGTTGTCTGTGAGGACTTTGGA	5883
Db	7092	GCAAACTCATATTTCGACAGGAAATATGATTAATCCTTAGTTGTCTGTGAGGACTTTGGA	7151
Qy	5884	GTCACTGTCTGTTGGACTTACTTCAACCAACTGATATGTCATGATGGGGTGGAGTTCAA	5943
Db	7152	GTCACTGTCTGTTGGACTTACTTCAACCAACTGATATGTCATGATGGGGTGGAGTTCAA	7211
Qy	5944	GATCAGGCAAGAGAAAAACATGTAAAAAGAAAGTAAATCTCCAACACTCACSSGGGTACATGGC	6003
Db	7212	GATCAGGCAAGAGAAAAACATGTAAAAAGAAAGTAAATCTCCAACACTCACCGGGTACATGGC	7271
Qy	6004	ACCTCTAGCCCTTACAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCAT	6063
Db	7272	ACCTCTAGCCCTTACAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCAT	7331
Qy	6064	ACTCGCCTGTGTAGCCCTATTTAATACCACCCCTCACTGGGTCCATGAGTCTCGGGCCAA	6123
Db	7332	ACTCGCCTGTGTAGCCCTATTTAATACCACCCCTCACTGGGTCCATGAGTCTCGGGCCAA	7391
Qy	6124	AACCTACTAATCTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	6183
Db	7392	AACCTACTAATCTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	7451
Qy	6184	GTACCTGAACAAATGGRACACTTCAGACAGAAATAAACACCACTTCGGTTTAGTAGGA	6243
Db	7452	GTACCTGAACAAATGGRACACTTCAGACAGAAATAAACACCACTTCGGTTTAGTAGGA	7511
Qy	6244	CCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTTAGC	6303
Db	7512	CCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTTAGC	7571
Qy	6304	AATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAATA	6363
Db	7572	AATACTACATACAAACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAATA	7631
Qy	6364	GTCTGCCCTACCCCTCAGGAATATTTTTGTCTGTGTACCTCAGCCTATCGTGTGTAAT	6423
Db	7632	GTCTGCCCTACCCCTCAGGAATATTTTTGTCTGTGTACCTCAGCCTATCGTGTGTAAT	7691
Qy	6424	GGCTCTTCAGAACTATGTGCTTCTCTCAATCTTTAGTGCCCCCYATGRCATCTACACT	6483
Db	7692	GGCTCTTCAGAACTATGTGCTTCTCTCAATCTTTAGTGCCCCCYATGRCATCTACACT	7751
Qy	6484	GAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAAAGAGTACCCATTCCT	6543
Db	7752	GAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAAAGAGTACCCATTCCT	7811
Qy	6544	CCTTTGTGTATAGGACAGAGTGCTAGGTGCACTAGGTACTGGCAATGGCGGTATACAA	6603
Db	7812	CCTTTGTGTATAGGACAGAGTGCTAGGTGCACTAGGTACTGGCAATGGCGGTATACAA	7871

QY	6604	ACCTCTACTCAGTTCCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAAACGGGTC	6663
DB	7872	ACCTCTACTCAGTTCCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAAACGGGTC	7931
QY	6664	GCCGACTCCCTGGTGCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCRA	6723
DB	7932	GCCGACTCCCTGGTGCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTTCAA	7991
QY	6724	AATCGAAGAGCTTTAGACTVGCCTAAACGGCTGABAGAGGGGGAAACCTGTGTTATTTTTAGGG	6783
DB	7992	AATCGAAGAGCTTTAGACTVGCCTAAACGGCTGAAAGAGGGGGAAACCTGTGTTATTTTTAGGG	8051
QY	6784	GAAGAACTGCTGTTATTATGTTTAACTCAATCCGGAATCGTCACTCAGAAAGGTTTAAAGAAAT	6843
DB	8052	GAAGAACTGCTGTTATTATGTTTAACTCAATCCGGAATCGTCACTCAGAAAGGTTTAAAGAAAT	8111
QY	6844	CSAGATCGAATACAACGTAKAGCAGAGAGCTTCGAAACACTGGACCCCTGGGGCCCTCCTC	6903
DB	8112	CGAGATCGAATCAACGTAGCAGAGAGGAGCTTCGAAACAACCTGGACCCCTGGGGCCCTCCTC	8171
QY	6904	AGCCRATGGATGCCCTGGATCTCCCTTCCTTAGGACCTCTAGCAGCTATAATATTGCTA	6963
DB	8172	AGCCAAATGGATGCCCTGGATCTCCCTTCCTTAGGACCTCTAGCAGCTATAATATTGCTA	8231
QY	6964	CTCCTCTTTGGACCCCTGTATCTTTTRACCTCCTTGTTTAACTTTGTCTCTCTCCAGAACTGAA	7023
DB	8232	CTCCTCTTTGGACCCCTGTATCTTTTAACTCTCTTGTTTAACTTTGTCTCTCCAGAACTGAA	8291
QY	7024	GCTGTRAAACTCAAAATGGAGCCCAAGATCGAGTCCAAGACTAAGATCTACCGCAGACCC	7083
DB	8292	GCTGTRAAACTCAAAATGGAGCCCAAGATCGAGTCCAAGACTAAGATCTACCGCAGACCC	8351
QY	7084	CTGACCGGGCTGYTAGCCCAAGATCTGATGTTAATGACATCAAAGGCAACCCCTCCTGAG	7143
DB	8352	CTGACCGGGCTGCTAGCCCAAGATCTGATGTTAATGACATCAAAGGCAACCCCTCCTGAG	8411
QY	7144	GAATCTCAGCTGCACAAACCTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGACGGT	7203
DB	8412	GAATCTCAGCTGCACAAACCTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGACGGT	8471
QY	7204	SGTCGGCCAACTCCCAACAGACCTTAGGTTTTCTGTTGATGAGTGGGGAC	7255
DB	8472	CGTCGGCCAACTCCCAACAGACCTTAGGTTTTCTGTTGATGAGTGGGGAC	8523

RESIT.T 4

```

US-09-854-867-21
; Sequence 21, Application US/09854867
; Patent No. 7014997
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC H
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: repeat_region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: herbv17
US-09-854-867-21

```

Query Match 81.8%; Score 6124.6; DB 5; Length 8523;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;  
Qv 606 GATGGGAAACGTTCCCGCAAGACAAAACGCCCTTAAGACGTATTCCTGGAAATTCGGGA 665

|||||  
1923 GATGGAAACGTTCCCGGACAGCAAAAGCGCCCTAAGACGTATTCTGGAGAAATGGGA 1982  
QY :  
666 MCAATTTGACCCCTCAGACACATAAGAAAGAAACGACTTATATTTCTTCTGAGTGCAGCGCTG 725  
Db :  
1983 CCAATTTGACCCCTCAGACACATAAGAAAGAAACGACTTATATTTCTTCTGAGTGCAGCGCTG 2042  
QY :  
726 GCACCTCCTGAGGGAAGTATAATTAATAACACCATCTTACAGCTAGACACTCTTTTGTAGAA 785  
Db :  
2043 GCACCTCCTGAGGGAAGTATAATTAATAACACCATCTTACAGCTAGACACTCTTTTGTAGAA 2102  
QY :  
786 AAGCAAAATGGAGTGAAGTGCCATAGTACAAACTTTCTTTTCAATTAAGAGACAACTCAC 845  
Db :  
2103 AAGCAAAATGGAGTGAAGTGCCATAGTACAAACTTTCTTTTCAATTAAGAGCAACTCAC 2162  
QY :  
846 AATTAATGAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATC 905  
Db :  
2163 AATTAATGAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATC 2222  
QY :  
906 CCAGCATCCCGACTCCTTCCCGCACTAATAAGGACCCCTTCAACCCAAATGCTCCAA 965  
Db :  
2223 CCAGCATCCCGACTCCTTCCCGCACTAATAAGGACCCCTTCAACCCAAATGCTCCAA 2282  
QY :  
966 AAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATTTATGA 1025  
Db :  
2283 AAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATTTATGA 2342  
QY :  
1026 CCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCCGCAGCCAGAGTGCATGTGCTTTTYY 1085  
Db :  
2343 CCCCT - CCAAGCAGTGGGAGGAGAAATTCGGCCCGCAGCCAGAGTGCATGTGCTTTTTC 2401  
QY :  
1086 TCTCCAGACTTAAAGCAAAATAAAAACAGACTTAGTAAATTTCTCAGATAAACCTTGATGG 1145  
Db :  
2402 TCTCCAGACTTAAAGCAAAATAAAAACAGACTTAGTAAATTTCTCAGATAAACCTTGATGG 2461  
QY :  
1146 CTATATGRTGTTTTACAGGGTTAGGACAAATTTCTTGATCTGCATGAGAGATATATA 1205  
Db :  
2462 CTATATGATGTTTTACAGGGTTAGGACAAATTTCTTGATCTGCATGAGAGATATA - A 2520  
QY :  
1206 TGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAGTGCACCACTAATCTGAGCCT 1265  
Db :  
2521 TGTCACTGCTAAATCAGACACTAACCCCAAATGAGAGAGTGCACCACTAATCTGAGCCT 2580  
QY :  
1266 GAGRTTTGGCGATCTGCTGATCTCAGTCAGGTCAA TGGATANGATGACAAAGAGG 1325  
Db :  
2581 GAGAGTTGGCGATCTGCTGATCTCAGTCAGGTCAA TG - ATAGGATGACAAAG - AGG 2637  
QY :  
1326 AAGANAATGATTTCCCAAGCCAGCARGCAGTTCCCAAGTCTASACCTCATTTGGGAC 1385  
Db :  
2638 AAGAGAAATGATTTCCCAAGCCAGCARGCAGTTCCCAAGTCTAGACCTCATTTGG - 2694  
QY :  
1386 ACAGAAATCAGTAACATGGGAGATTTGGTCTCAGACATTTGCTAACTTTGTGCTASAA 1445  
Db :  
2695 ACACAGAAATCAGACAT - GGAGATTGGTCTCAGACATTTGCTAACTTTGTGCTAGAA 2753  
QY :  
1446 GGACTAAGGAAACTAGAGAAATCTAYGAATTAATCAATGATGTCCACCATACACA 1505  
Db :  
2754 GGACTAAGGAAACTAGGAAG - AAGTCTATGAATTAATCAATGATGTCCACCATACACA 2812  
QY :  
1506 GGGAGGGAAGAAATCTACTGCTTCTGAGAGACTAAGGAGGCAATTTAGGAAGC 1565  
Db :  
2813 - GGGAGGGAAGAAATCTACTGCTTCTGAGAGACTAAGGAGGCAATTTAGGAAGC 2871  
QY :  
1566 GTGCTCTCTGTCACCTGACTCTTCTGAAAGGCCAACTAATCTTAAGCGTAAGTTATCA 1625  
Db :  
2872 GTGCTCTCTGTCACCTGACTCTTCTGAAAGGCCAACTAATCTTAAGCGTAAGTTATCA 2931  
QY :  
1626 CTCAGTCAGCTGAGACATTAAG - AAAAACTTCAAAAGTCTGCGGTAGGCCCGGAGCAAA 1684  
Db :  
2932 CTCAGTCAGCTGAGACATTAAGAAAAAACTTCAAAAGTCTGCGGTAGGCCCGGAGCAAA 2991  
QY :  
1685 ACTTAGAAACCTATTGAACCTTGGCAACTGCGTGTATTAATATAGAGATCAGGAGGAGC 1744  
|||||

Db :  
2992 ACTTAGAAACCTTATGAACCTTGGCAACCTCGGTTTTTTTATAATAGAGATCAGGAGGAGC 3051  
QY :  
1745 AGGCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTTAGTCATGACCCCTCAGG 1804  
Db :  
3052 AGGCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTTAGTCATGACCCCTCAGG 3111  
QY :  
1805 CAACTGGACCTTTGAGAGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAATAGG 1864  
Db :  
3112 CAACTGGACCTTTGAGAGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAATAGG 3171  
QY :  
1865 GCTTGTCTCCAGTGCAGTCTCAAGAGGACATTTTAAAAAAGATTTGTCCAAGTAGAAGTAAG 1924  
Db :  
3172 GCTTGTCTCCAGTGCAGTCTCAAGAGGACATTTTAAAAAAGATTTGTCCAAGTAGAAGTAAG 3231  
QY :  
1925 CCGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 1984  
Db :  
3232 CCGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 3291  
QY :  
1985 GGACAAAGTCTTTTGGAGTCAAGAACCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG 2044  
Db :  
3292 GGACAAAGTCTCTGAGTCAAGAACCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG 3351  
QY :  
2045 TGCCTGGGGCAAGCGCCATCCCATCCATCACCTCACAGAGCCCTGGGTATGCTTGACC 2104  
Db :  
3352 TGCCTGGGGCAAGCGCCATCCCATCCATCACCTCACAGAGCCCTGGGTATGCTTGACC 3411  
QY :  
2105 ATTGAGGCCAAGGAAGTGTCTCTGAGCACTGCTGGTGGGTCTTCTTAGTCTTACTCTTC 2164  
Db :  
3412 ATTGAGGCCAAGG - AGGTTGTCTCTGAGCACTGCTGGCGGTCTTCTTAGTCTTACTCTTC 3470  
QY :  
2165 TGTCCCGGACAACTGCTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGGGCA 2224  
Db :  
3471 TGTCCCGGACAACTGCTCTCCAGATCTGTCACTA - TCTGAGGGGTCC - TAAGCGGGCA 3528  
QY :  
2225 GTCACTAGATACTTTTCCAGCCACTAAAGTTATGAATCGGGGAGCTTTATTTCTTTTTCAC 2284  
Db :  
3529 GTCACTAGATAC - TTTCTCCAGCCACTAAAGTTATG - ACTGGGAGCTTTATTTCTTTTTCAC 3586  
QY :  
2285 ATGCTTTTCTAAATATGCTTGAAGCCCACTACCTGTTTAGGGAGAGACATTTCTAGCAA 2344  
Db :  
3587 ATGCTTTTCTAAATATGCTTGAAGCCCACTACCTGTTTAGGGAGAGACATTTCTAGCAA 3646  
QY :  
2345 AAGCAGGGCCCATTTATACACCTGAACATAGAGAGGAACACCGTTTGTGTGTCNCCCTG 2404  
Db :  
3647 AAGCAGGGCCCATTTATACACCTGAACATAGAGAGGAACACCGTTTGTGTG - CCCCTG 3705  
QY :  
2405 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGACATATGAGCAGGCCAAA 2464  
Db :  
3706 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGACATATGAGCAGG - CAAA 3764  
QY :  
2465 GAATGCCGCTCCTGTTCAAGTTAAACTAAAGATTTCCAAGATTTGTTAAGGACTTAAAGC 2524  
Db :  
3765 GAATGCCGCTCCTGTTCAAGTTAAACTAAAGATTTCCAACCTCTTCCCTTACCAAAGGCA 3824  
QY :  
2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAAAGC 2584  
Db :  
3825 GTACCCCTCAGACCCCAAGGCCCAACAGGACTTTCCAAAGATTTGTTAAGGACTTAAAGC 3884  
QY :  
2585 CCAAGGCTTAGTAAAAACCATGATAACTCTCCCTGCAAGTAAATTCCTGATGATTTAGGAGG 2644  
Db :  
3885 CCAAGGCTTAGTAAAAACCATGATAACTCTCCCTGCAAGTAACTCC - - - - - AATTTTAGGAG 3937  
QY :  
2645 CACAGAAACCAAGTGAACAGTGGGGTTAGTGAAGATTTCTCAGGATTTATCAATGAGGCG 2704  
Db :  
3938 TACAGAAACCAACAGACAGTGA - GGTTAGTGAAGATTTCTCAGGATTTATCAAT - GAGGC 3995  
QY :  
2705 GGTGTCTCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCTCAATACCAGA 2764  
Db :  
3996 TGTGTTCTCTTAGCAGCAGCTGTACCTAGCCCTTATCTGTCTTTCTCCAAATACCAGA 4055  
QY :  
2765 GGAAGCAGAGTGGTTTACASTCTCGGACCTTMMAGATGCTTTCTTCTGTCATCCCTGTACA 2824  
Db :  
4056 GGAAGCAGAGTGGTTTACAGTCTTGGACCTTCAGGATGCTTCTTCTGTCATCCCTGTACA 4115  
|||||



[illegible]

Db	7332	ACTGGCCTGGTAGAGCCTATTTTAATACCAACCTTCACCTGGGCTCCATGAGAGCTCTCGGCCCAA	7391
Qy	6124	AACCCCTACTAACTGTTGGATATGCTCTCCCTCCCTGAACTTCARGCCCATATGTTTTCAATCCCT	6183
Db	7392	AACTCTACTTAACTGTTGGATATGCTCTCCCTGAACTTCAGGCCCATATGTTTCAATCCCT	7451
Qy	6184	GTACTCTGAAACAATGGAAACAACCTTCAGCACAGAAAAAACAACAACCTTCGGTTTTAGTAGGA	6243
Db	7452	GTACTCTGAAACAATGGAAACAACCTTCAGCACAGAAAAAACAACAACCTTCGGTTTTAGTAGGA	7511
Qy	6244	CCCTCTGTTTCCAAATSTGGAAATACCCCATACCTCAACCTCACTGCTGTGTAAAAATTAGC	6303
Db	7512	CCCTCTGTTTCCAAATSTGGAAATACCCCATACCTCAACCTCACTGCTGTGTAAAAATTAGC	7571
Qy	6304	AATACTACATACACAACCAACCTCCCAATCAGTCAGGTGGGTAACTCCTCCCAACACAAAATA	6363
Db	7572	AATACTACATACACAACCAACCTCCCAATCAGTCAGGTGGGTAACTCCTCCCAACACAAAATA	7631
Qy	6364	GTCTGCTTACCTTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCGTGTGTTTGAAT	6423
Db	7632	GTCTGCTTACCTTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCGTGTGTTTGAAT	7691
Qy	6424	GGCTCTTCAGAACTATATGTGCTTCCTCTCATCTTCTTAGTGCCCCCATGBCCAATCTACACT	6483
Db	7692	GGCTCTTCAGAACTATATGTGCTTCCTCTCATCTTCTTAGTGCCCCCATGBCCAATCTACACT	7751
Qy	6484	GAACAAGATTATACAGTTATGTCTATCTAAGACCCCGCAACAAGAGAGTACCCATTCTT	6543
Db	7752	GAACAAGATTATACAGTTATGTCTATCTAAGACCCCGCAACAAGAGTACCCATTCTT	7811
Qy	6544	CCTTTTGTTATAGGACGAGAGTGCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATCAACA	6603
Db	7812	CCTTTTGTTATAGGACGAGAGTGCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATCAACA	7871
Qy	6604	ACCTCTACTCAGTTCTACTACAAAATCTATCTCAAGAACTAAATGGGGACATCGAAACGGGTC	6663
Db	7872	ACCTCTACTCAGTTCTACTACAAAATCTATCTCAAGAACTAAATGGGGACATCGAAACGGGTC	7931
Qy	6664	GCGACTCCCTGTCTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTCRA	6723
Db	7932	GCGACTCCCTGTCTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTCNA	7991
Qy	6724	AATCGAAGAGCTTTAGACTYGTCTAACCGCTGARAGAGGGGAACTGTGTTATTTTTAGGG	6783
Db	7992	AATCGAAGAGCTTTAGACTYGTCTAACCGCTGARAGAGGGGAACTGTGTTATTTTTAGGG	8051
Qy	6784	GAAGATGCTGTTTATTTATGTTAATCAATCGGAAATCGTCACTGAGAGAAAGTTAAGAAAT	6843
Db	8052	GAAGATGCTGTTTATTTATGTTAATCAATCGGAAATCGTCACTGAGAGAAAGTTAAGAAAT	8111
Qy	6844	CSAGATCGAATACAACTAKAGCAGARGAGCTTCGAAACAACCTGGACCTCGGGGCTCTC	6903
Db	8112	CGAGATCGAATACAACTGAGCAGAGAGCTTCGAAACAACCTGGACCTCGGGGCTCTC	8171
Qy	6904	AGCCRATGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA	6963
Db	8172	AGCCRATGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA	8231
Qy	6964	CTCCTCTTTGGAGCCCTGTATCTTTTTRACCTCTCTGTTTAACTTTGTCTCTTCGAGATCGAA	7023
Db	8232	CTCCTCTTTGGAGCCCTGTATCTTTTAACTCTCTGTTTAACTTTGTCTCTTCGAGATCGAA	8291
Qy	7024	GCTGTAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGACCC	7083
Db	8292	GCTGTAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGACCC	8351
Qy	7084	CTGGACCGGCTCTGTAGCCCAACGATCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAG	7143
Db	8352	CTGGACCGGCTCTGTAGCCCAACGATCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAG	8411
Qy	7144	GAATCTCAGCTGCAACAACCTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGACGGT	7203
Db	8412	GAATCTCAGCTGCAACAACCTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGACGGT	8471









```
Db 16205 --ATGGCCCCCGGCGTCATATTTCTCTTACTGTGCTTACCCCGCTTTCACATCA 16148
Qy 5639 CTGACCCCTCCATGCGGTGTATGACCACTAGTCCCTTACCMAGAGTTTCTATGGA 5698
Db 16147 CCTCACCCCTCCATGCGGTGTATGACCACTAGTCCCTTACCAAGAGCTTCTATGGA 16088
Qy 5699 GAATGAGCGCTCCCGGAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACC 5758
Db 16087 GAATGCGGCTTCCAGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACC 16028
Qy 5759 CCACCTTCACTGCCACCACTATGCCCGCAACTGC 5796
Db 16027 CCACCTTCACTGCCACCACTATGCCCGCAACTGC 15990

RESULT 6
US-09-949-016-12249/c
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

Query Match 40.0%; Score 2994.2; DB 3; Length 77997;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;

Qy 1812 ACTTTGGAGGCTCTGAAAGGGAAAGCTGGGCAAAATTGAATGCCCTAATAGGGCTTGCT 1871
Db 19935 ACTTTGGAGGCTCTGAAAGGGAAAGCTGGGCAAAATTGAATGCCCTAATAGGGCTTGCT 19876

Qy 1872 TCCAGTGGCGGTCTACAGGACACTTTAAAGAGATTGTCCAGTAGAAGTAAAGCCGCC 1931
Db 19875 TCCAGTGGCGGTCTACAGGACACTTTAAAGAGATTGTCTGAATAGAAATAAGCTGCC 19816

Qy 1932 TTCGTCCATGCCCTTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCGGAGCAAA 1991
Db 19815 -TCATCCATGCCCTCATGTCAAGGGAATCACTGGAAGGCCCACTGCCCGGAGTAA 19757

Qy 1992 GGTCTTTTGGTCTGAGAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGGTGCTGG 2051
Db 19756 GGTCTTTTGGTCTGAGAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGGTGCTGG 19697

Qy 2052 GGCAGGCGCATCCCATGCCATCACCCCTCAGAGAGCCCTGGGTATGCTTGACCATGAGG 2111
Db 19696 GGCAGGCGCATCCCAT-----GCCCGAGGTATGCTTGACCATGAGG 19654

Qy 2112 GCCAGGAAGGT---TGCTCTCGGACACTGGTGGGTCTTCTTAGTCTTACTTCTGTC 2168
Db 19653 GCCAGGAAGGTAACTGCTCTCCTGAACTGGCATGGCATGGCCTTCTCAGTCTTAGTCTGCTG 19594
```

```
Qy 2169 CCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCNTAAGACGGGCGAGTCA 2228
Db 19593 CTGGACAAACCGTCTCTCCAGATCTGTCACTACCC---GAGGGGTCCTAGGATAGGCGAGTCA 19537

Qy 2229 CTAGATACTTTTCCAGGACCACTAACTATGAATCGGAGCTTTATTTCTTTTCACATGC 2288
Db 19536 CTAGATAC-TTCTCCAGGACCACTAAGTTGTG-ACTGGGAACTTCAITCTTTTCACTTGC 19479

Qy 2289 TTTTCTAAATATGCTTTGAAAGCCCACTACCTTTGTTTAGGGAGAGACATTTCTAGCAAAAGC 2348
Db 19478 CTTTCTAAATATGCTTTGAAAGCCCACTACCTTTGTTTAGGGAGATACATTTCTAGCAAAAGC 19419

Qy 2349 AGGGGCCATATACACTGAAACATAGGAGAGAAACACCGCTTTGTTGTTGTTGTTGTTGTTG 2408
Db 19418 AGGGGCCATATATACACTGAAACATAGGAGAGAAACACCGTTTGTCT--GTCCCTCTACTT 19361

Qy 2409 AGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAAATATGACGACGAGCCAAAGAAT 2468
Db 19360 AGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAAATATGATGAG-GGAAGAAT 19302

Qy 2469 GCCCGTCTGTTCAAAGTTAAACTTAAAGGATCCACTTCTTCCCTTCCCTTCCCTTCCCTTCCCTT 2528
Db 19301 GCCCGTCTGTTCAAAGTTAAACTTAAAGGATTCGCGCTCTTCTTCCCTTCCCTTCCCTTCCCTT 19242

Qy 2529 CCCTCAGACCCCAAGGCCCAACAGGATTCCTAAAGATTTGTTAAGGACTTAAAGCCCAA 2588
Db 19241 CCTCTTAGACCCGAGGCCCAACAGGACTTCCAAAGATTTGTTAAGGACCTTAAAGCCCAA 19182

Qy 2589 GCTTAGTAAACCACTGATTAATCTCCCTGACGATTAATCCGTAGTGGATTGAGGAGGCACA 2648
Db 19181 GCTTAGTAAACCACTGATTAATCTCCCTGACGATTAATCCGTAGTGGATTGAGGAGGCACA 19129

Qy 2649 GAAACCCAGTGGACAGTGGAGGGTTAGTCAAGATCTCAGGATTAATCAATGAGGAGCGGTT 2708
Db 19128 GAAACCCAGTGGACAGTGGAGGGTTAGTCAAGATCTCAGGATTAATCAATGAGGAGCGGTT 19071

Qy 2709 GTCTCTTTTATACCCAGTGTACTAGCCCT---TATACTGTGTGMYTTTCCAAATACCAGAG 2765
Db 19070 GTTCTCTTATACCCAGTGTACTAGCCCT---TATACTGTGTGMYTTTCCAAATACCAGAG 19011

Qy 2766 GAAGCAGAGTGGTTTACASTGCTGGACCTTMAAGATGCTTCTTCTGATCCCTGTACAT 2825
Db 19010 GAAGCAGAGTGGTTTACAGTCTCTGACCTTAAAGATGCTTCTTCTGATCCCTGTACAT 18951

Qy 2826 CCTGACTCTCAATTTCTTGTGCTTTGAGATGACTTTCAAAACCCARCATCTCAACTCACC 2885
Db 18950 CCTGACTCTCAATTTCTTGTGCTTTGAGATGCTTTGAAATCTAATGTCTCAACACACC 18891

Qy 2886 TGGACTTTTATACCCCAAGGTTTCAAGGATAGYCCCCATCTATTGTCGAGGAGCATTTAGCC 2945
Db 18890 TGGACTTTTATACCCCAAGGTTTCAAGGATAGYCCCCATCTATTGTCGAGGAGCATTTAGCC 18831

Qy 2946 CAAGACTTGAGTCATYMTCACTGAGACCTTGTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG 3005
Db 18830 CAAGACTTGAGTCATYMTCACTGAGACCTTGTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG 18771

Qy 3006 CTTTTRGCGVCCYRITTCAGAAACCTTTGSCCATCAAGCCACCCAGCRCTCTTMAATTTTC 3065
Db 18770 CTTTTRGCGVCCYRITTCAGAAACCTTTGSCCATCAAGCCACCCAGCRCTCTTMAATTTTC 18711

Qy 3066 CTCGCGYACCTGTGGCTACAWGGTTTCCAAACASARAGCTCARCTCTGCTCAGACGAGGTT 3125
Db 18710 ATGGCCACCTGTGGCTACAWGGTTTCCAAACASARAGCTCARCTCTGCTCAGACGAGGTT 18651

Qy 3126 AATATCTTAGGCTTAARATTTATCCAAAGGACACAGGGCCCTCAGTGAGGAAATATATCCAG 3185
Db 18650 AATATCTTAGGCTTAARATTTATCCAAAGGACACAGGGCCCTCAGTGAGGAAATATATCCAG 18591

Qy 3186 CCTATCTGGCTTATCTCATCYGAAACCCCTTAAAGCACTAAGGRRTTCTTGGCPTA 3245
Db 18590 CCTATCTGGCTTATCTCATCYGAAACCCCTTAAAGCACTAAGGRRTTCTTGGCPTA 18531

Qy 3246 AVAGGYTTCTGCGGAATGATGTTCCCGAGGTTGCGCAATAGCCAGGYCATTAATATAC 3305
```

[illegible]

```
QY 5460 CTGCCAAGGAACCTGAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAG 5519
D 16329 CCGCAAGGAACCAAAATTCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAG 16270
QY 5520 GATTGGCGCTGCTCTTCAAAACAACCAAGGAGAAAGTAATCAATAATCAATAATCCCC 5579
D 16269 GATCTGCACCTGCTCTTCAAGCGACAACTGCGAGGAAGTAATCGTAGAGCTCC 16210
QY 5580 CATGSCCTCCCTTATCATATTTCTCTKTASTGTTSTTYA-CCCTSTTTCACTCA 5638
D 16209 --ATGGGCCCCCGCTCATATTTCTTTACTTGTGTCTTACCCCTTTCACTATCA 16152
QY 5639 CTGCACCCCTCCATCGCGCTGATGACCACTAGCTCCCTTACCAAGAGTTTCTATGGA 5698
D 16151 CCTCACCCCTCCATGCGCTATACCTACAGTAGCTCCCTTACCAAGAGCTTCTATGGA 16092
QY 5699 GAATGACGCGTCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCC 5758
D 16091 GAATGCGCTTCCAGAAATATTGATGCCCATTTGATGCCCATTTGATAGGAGTTTCTTAAGGGAACCC 16032
QY 5759 CCACCTTCACTGCCCAACCCATATGCGCCGCAACTGC 5796
D 16031 CCACCTTCACTGCCCAACCCATATGCGCCGCAACTTC 15994
```

## RESULT 7

```
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Metberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3
```

```
Query Match 38.6%; Score 2888.6; DB 3; Length 2946;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2872; Conservative 49; Mismatches 3; Indels 1; Gaps 1;

QY 4658 CTGCTVATCGCAAGCTCTTCAAGGAAACAARAACAGGCCATTACCTGPARAARAC 4717
D 6 CTGCTTATCGCAAGCTCTTCAAGGAAACAARAACAGGCCATTACCTGGAAGAC 65
QY 4718 TGGCACTGATTTTACCACAGCCCAACCTCAGGGATTTTCACTAGTCTGGG 4777
D 66 TGGCACTGATTTTACCACAGCCCAACCTCAGGGATTTTCACTAGTCTGGG 125
QY 4778 TATATCTTTTACGGGTTGGGAGAGCCCTTCCCTGTAGGACAGAAAGGCCAAGAGG 4837
D 126 TATATCTTTTACGGGTTGGGAGAGCCCTTCCCTGTAGGACAGAAAGGCCAAGAGG 185
QY 4838 TAATAAGGCACTAGTTTCAATGAATAATTCAGACTTCGGACTTCCCGAGGCTTACAGA 4897
D 186 TAATAAGGCACTAGTTTCAATGAATAATTCAGACTTCGGACTTCCCGAGGCTTACAGA 245
QY 4898 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCAGGGCGTTAGGTA 4957
```

```
Db 246 GTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTA 305
QY 4958 TACGATATCACTTACACTGCGCTTGAAGGCCACAGTCTCAGGGAAGGTGAGAAATGA 5017
D 306 TACGATATCACTTACACTGCGCTTGAAGGCCACAGTCTCAGGGAAGGTGAGAAATGA 365
QY 5018 ATGAAAVACTCAAGAGGACATCTAAAAGCAAAACCCAGGAAACCCACTCACAATGGGCTG 5077
D 366 ATGAAACACTCAAGAGGACATCTAAAAGCAAAACCCAGGAAACCCACTCACAATGGGCTG 425
QY 5078 YTCGTGTGCTATAGCCCTTAAAAGATCTGCAACTTTCCCAAAAAGCAGGACTTAGCC 5137
D 426 CTCGTGTGCTATAGCCCTTAAAAGATCTGCAACTTTCCCAAAAAGCAGGACTTAGCC 485
QY 5138 CATAAGAAATGCTGTATGGAAGGCCCTTCAATAACCAATGACCTTGTGCTTGACCACAAGAC 5197
D 486 CATAAGAAATGCTGTATGGAAGGCCCTTCAATAACCAATGACCTTGTGCTTGACCACAAGAC 545
QY 5198 AGCCAACTTATGTTGAGACATCACTCTCTTAGCCAAATATCAACAAGTTCTTAAACAATT 5257
D 546 AGCCAACTTATGTTGAGACATCACTCTCTTAGCCAAATATCAACAAGTTCTTAAACAATT 605
QY 5258 ACAAGGAACCTATCCCTGAGAGGGAAGAAAGAACTATTCCACCCWGTGACATGTTATT 5317
D 606 ACAAGGAACCTATCCCTGAGAGGGAAGAAAGAACTATTCCACCCWGTGACATGTTATT 665
QY 5318 AGTCAAGTCCCTTCYCTCTTAATTTCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCC 5377
D 666 AGTCAAGTCCCTTCYCTCTTAATTTCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCC 725
QY 5378 AGTCATTTTATYACCCCAACTCGCGTTAAAGTGGTGGAGTGGAGTCTTTGGATACATCA 5437
D 726 AGTCATTTTATYACCCCAACTCGCGTTAAAGTGGTGGAGTGGAGTCTTTGGATACATCA 785
QY 5438 CACTTGAGTCAAAATCTGAGTACTGCTCAAGGAACCTGAAAATCCAGGAGACAAACGCTAG 5497
D 786 CACTTGAGTCAAAATCTGAGTACTGCTCAAGGAACCTGAAAATCCAGGAGACAAACGCTAG 845
QY 5498 CTATTCCTGTGAACCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAACAACAGGAGAAA 5557
D 846 CTATTCCTGTGAACCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAACAACAGGAGAAA 905
QY 5558 GTAACTAAAATCATAAATCCCATGSCCTCCCTTATCATATTTTCTCTKTASTGTTTS 5617
D 906 GTAACTAAAATCATAAAT-CCCCATGSCCTCCCTTATCATATTTTCTCTKTASTGTTTC 964
QY 5618 TTTTACCCCTTTTCACTCTCACTGACCCCTCCATGCGCTGTATGACCAAGTAGCTCCC 5677
D 965 TTTTACCCCTTTTCACTCTCACTGACCCCTCCATGCGCTGTATGACCAAGTAGCTCCC 1024
QY 5678 CTYACCMAGAGTTTCTATGAGAGATGCGGTCCTCCGGAATATTTGATGCCCCCATCGTATA 5737
D 1025 CTYACCMAGAGTTTCTATGAGAGATGCGGTCCTCCGGAATATTTGATGCCCCCATCGTATA 1084
QY 5738 GGAGTCTTTSTAAGGAAACCCCACTTCACTGCCCAACCCATATGATGCCCCCAACTGCT 5797
D 1085 GGAGTCTTTSTAAGGAAACCCCACTTCACTGCCCAACCCATATGATGCCCCCAACTGCT 1144
QY 5798 ATCACTCTGCACTCTTTTGCATGCAATACTCATTTATTTGGAACAGGAAAAATGATTA 5857
D 1145 ATCACTCTGCACTCTTTTGCATGCAATACTCATTTATTTGGAACAGGAAAAATGATTA 1204
QY 5858 ATCTCTAGTTGCTCTGAGGACTTGGAGTCACTGCTGTGTTGGACTTCTTACCACCAACTG 5917
D 1205 ATCTCTAGTTGCTCTGAGGACTTGGAGTCACTGCTGTGTTGGACTTCTTACCACCAACTG 1264
QY 5918 GTATGCTCATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAGAAAGATTA 5977
D 1265 GTATGCTCATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAGAAAGATTA 1324
QY 5978 TCTCCCAACTCAGCCGGGTACATGGCACTCTTAGCCCTTACAAAGGACTAGATCTCTCAA 6037
```

Db 1325 TCTCCAACTACACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAA 1384  
 QY 6038 AACTACATGAAGAACCTCCGTACCCATACATCGCTGCTAGAGCTATTTAATACCAACCTCA 6097  
 Db 1385 AACTACATGAAGAACCTCCGTACCCATACATCGCTGCTAGAGCTATTTAATACCAACCTCA 1444  
 QY 6098 CTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCGA 6157  
 Db 1445 CTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCGA 1504  
 QY 6158 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAATGGAACAACTTTCAGCACAGAAA 6217  
 Db 1505 ACTTCARGCCATATGTTTCAATCCCTGTACCTGAACAAATGGAACAACTTTCAGCACAGAAA 1564  
 QY 6218 TAAACACCACTTCCTGTTTGTAGGACCTCTGTTTCCAAATGTTGGAATTAACCCATACCT 6277  
 Db 1565 TAAACACCACTTCCTGTTTGTAGGACCTCTGTTTCCAAATGTTGGAATTAACCCATACCT 1624  
 QY 6278 CAAACCTCAGCTGTGTAATATTTAGCAATACCTACATACACCAACCACTCCCAATGCATCA 6337  
 Db 1625 CAAACCTCAGCTGTGTAATATTTAGCAATACCTACATACACCAACCACTCCCAATGCATCA 1684  
 QY 6338 GGTGGGTAACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTG 6397  
 Db 1685 GGTGGGTAACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTG 1744  
 QY 6398 GTACCTCAGCCTATCGTGTGTTTGAATGGCTCTTCAGAAATCTATGCTTCTCTCATCT 6457  
 Db 1745 GTACCTCAGCCTATCGTGTGTTTGAATGGCTCTTCAGAAATCTATGCTTCTCTCATCT 1804  
 QY 6458 TAGTGGCCCCYATGRCATCTACACTGAACAGATTTATACAGTTATGTCATATCTTAGC 6517  
 Db 1805 TAGTGGCCCCYATGRCATCTACACTGAACAGATTTATACAGTTATGTCATATCTTAGC 1864  
 QY 6518 CCGGCAACAAAGAGTACCATCTCTCTTTGTTTATAGGAGCAGGAGTCTAGTGCAC 6577  
 Db 1865 CCGGCAACAAAGAGTACCATCTCTCTTTGTTTATAGGAGCAGGAGTCTAGTGCAC 1924  
 QY 6578 TAGGTACTGGCAATGCGCGTATCACAACTCTACTCAGTTCTACTACAAACTATCTCAAG 6637  
 Db 1925 TAGGTACTGGCAATGCGCGTATCACAACTCTACTCAGTTCTACTACAAACTATCTCAAG 1984  
 QY 6638 AACTAAATGGGACATGGAACGGGTCCGCACTCCCTGCTCAGCTTGGAGATCACTTA 6697  
 Db 1985 AACTAAATGGGACATGGAACGGGTCCGCACTCCCTGCTCAGCTTGGAGATCACTTA 2044  
 QY 6698 ACTCCCTAGCAGCAGTGTCTTTCRAAATCGAAGAGCTTTAGACTGCTTAAACCGCTGARA 6757  
 Db 2045 ACTCCCTAGCAGCAGTGTCTTTCRAAATCGAAGAGCTTTAGACTGCTTAAACCGCTGARA 2104  
 QY 6758 GAGGGGAACTGTTTATTTTATAGGGGAAAGATGCTGTTTATTATGTTAATCAATCCGAA 6817  
 Db 2105 GAGGGGAACTGTTTATTTTATAGGGGAAAGATGCTGTTTATTATGTTAATCAATCCGAA 2164  
 QY 6818 TGTCTACTGAGAAAGTGAAGAAATTCAGATCGAATACAACTAGAGCAGAGAGCTTC 6877  
 Db 2165 TGTCTACTGAGAAAGTGAAGAAATTCAGATCGAATACAACTAGAGCAGAGAGCTTC 2224  
 QY 6878 GAAACACTGGACCTCGGGGCTCTCAGCCATGAGTGCCTGATCTTCCCTCTCTTAG 6937  
 Db 2225 GAAACACTGGACCTCGGGGCTCTCAGCCATGAGTGCCTGATCTTCCCTCTCTTAG 2284  
 QY 6938 GACCTCTAGCAGCTAATATTTGCTACTCTCTTTTGGACCTGTATCTTTTACCTCCTTG 6997  
 Db 2285 GACCTCTAGCAGCTAATATTTGCTACTCTCTTTTGGACCTGTATCTTTTAACTCCTTG 2344  
 QY 6998 TTAACCTTGTCTCTCCAGATCGAAGCTGTAACTACAAATGGAAGCCCAAGATGCACT 7057  
 Db 2345 TTAACCTTGTCTCTCCAGATCGAAGCTGTAACTACAAATGGAAGCCCAAGATGCACT 2404  
 QY 7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTA 7117  
 Db 2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTA 2464

QY 7118 ATGACATCAAAAGGCACCCCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCA 7177  
 Db 2465 ATGACATCAAAAGGCACCCCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCA 2524  
 QY 7178 ATTACAGAGGAAGCAGATTAGAGCGGCTGTCGGCCAACTCCCCAAACAGACACTTAGGTTT 7237  
 Db 2525 ATTACAGAGGAAGCAGATTAGAGCGGCTGTCGGCCAACTCCCCAAACAGACACTTAGGTTT 2584  
 QY 7238 CTTGTTGAGATGCGGGGACTGAGAGACAGGACTAGCTGATTTCTTAGGCTGATTAAGAAAT 7297  
 Db 2585 CTTGTTGAGATGCGGGGACTGAGAGACAGGACTAGCTGATTTCTTAGGCTGATTAAGAAAT 2644  
 QY 7298 CCYTAAGCCTAGSTGGGAAGGTGACACATCCACCTTTAAACACCGGGCTTGCACCTTAG 7357  
 Db 2645 CCCTAAGCCTAGCTGGGAAGGTGACACATCCACCTTTAAACACCGGGCTTGCACCTTAG 2704  
 QY 7358 YTCACACCTGACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAAGACAGGAGTTAA 7417  
 Db 2705 CTCACACCTGACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAAGACAGGAGTTAA 2764  
 QY 7418 AGAATATGCCAATCATYTTATTGCMTGAGACACAGCAGAGGGAACAATGATCGGATATA 7477  
 Db 2765 AGAATATGCCAATCATCTATTGCTGAGACACAGCAGAGGGAACAATGATCGGATATA 2824  
 QY 7478 AACCACAGTCTCGAGCGGCAACCGGCAACCCCTTTGGTCCCTCTTGTATGGGA 7537  
 Db 2825 AACCACAGTCTCGAGCGGCAACCGGCAACCCCTTTGGTCCCTCTTGTATGGGA 2884  
 QY 7538 GCTCTGTTTTCATGCTATTCTCACTCTAATTAATCTTGCARCTGCR 7582  
 Db 2885 GCTCTGTTTTCATGCTATTCTCACTCTAATTAATCTTGCARCTGCA 2929

RESULT 8  
 US-09-949-016-13002/c  
 ; Sequence 13002, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; PRIORITY FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13002  
 ; LENGTH: 168394  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(168394)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-13002

Query Match 34.9%; Score 2615.4; DB 3; Length 168394;  
 Best Local Similarity 85.1%; Pred. No. 0;  
 Matches 3133; Conservative 134; Mismatches 355; Indels 58; Gaps 33;  
 QY 1012 TTCCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCAGCCAGAGTG 1071  
 Db 29005 TCCCCCAAGGCAAAACACCCCTTAAATGTATCTGGAGAAATTCGGCTTGTACAGTG 28946  
 QY 1072 CATGTGCTTTTCTTCTCCAGACTTAAAGCAAAATAAAGACACTTAGTAAATTTCTCA 1131

Db	28945	TATGTACCTTTTCCCTGTGACAGCTTGAAGCAAAATTAATAAGACCTAGGTAAATTTCTCA	28886
Qy	1132	GATAAYCCTGATGCTATATGRTGTTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACA	1191
Db	28885	GATAAACCTGTGTGCTATATGATGCTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACA	28826
Qy	1192	TGGAGAGATATATGTCACCTGCTAAATACAGACACTAACCCCAATCAGAGAGTGCAC	1251
Db	28825	TGGAGAGATATA-ATGTTACTGCTAGATCAGACACTAATCCCAATCAGAGAGTGTCTG	28767
Qy	1252	CATAACTGACGCTGAGGTTTGGCGATCTCTGCTATCTCAGTCAGTCGAATGGATANGG	1311
Db	28766	CATAACTGACGCTGAGAGTTTGGTGATCTCTGCTATCTCAGTCAGTCGAATG-ATAGG	28709
Qy	1312	ATGACAAAGAGAGAAAGANAATGATTTCCCAACAGGCGACGACAGTCTCCAGCTCTASA	1371
Db	28708	ATGACAAAG-AGNAAGAGACAAATTTCCCAACAGGCGACGAGTCTCCAGTGTAGA	28650
Qy	1372	CCCTCATTTGGGACACAGAAATCAGTTAAACATGGAGATTTGGTCTGCAGACATTTGCTAA	1431
Db	28649	CCCTCATTTGGG-ACACAGAAATCAGACATGGAATTTGGTCCACAGACATTTGCTAA	28595
Qy	1432	CTTGTGTGCTASAGGACTAAGGAAACTAGSAGAAATCTAYGAATTAATCTCAATGATG	1491
Db	28594	CTTGTGTGCTA-AAGGACTAAGGAAACTTAGAAG-AGGCTATAGATTAATCAGTATG	28537
Qy	1492	TCCACCATACACAGGGGAGGAGAAATCTCTAGTCTTTCTGGAGAGACTAAGGGA	1551
Db	28536	TTCACTATATACCA-AGGAAGGAGAAATCTCTAGTCTTTCTGGAGAGACTAAGGGA	28478
Qy	1552	GGCATTTGAGGAAGCTGCTCTCTGTCTCAGCTGACTCTCTGAAGGCCCACTAATCTTAA	1611
Db	28477	GGCATTTGAGGAGC-AAACTCTCTGTCTCAGCTGACTCTGTGAAGGCCCACTAATCTTAA	28419
Qy	1612	GGCTAAGTTTATCACTCAGTCAGTCGACGATAG-AAAAAACTTCAAAAGTCTGCCGT	1670
Db	28418	GGTTAAGTTTATCACTCAGTCAGTCGACGATAGGAAATTAAGTTTCAACCTT	28359
Qy	1671	AGGCGGAGAGCAAACTTAGAAACCTTATGAACTTGGCAACVTCGGTTTTTATATAG	1730
Db	28358	AGGCGGAGAGCAAACTTAGAAACCTTATGAACTTGGCAACVTCGGTTTTTATATAG	28299
Qy	1731	AGATCAGGAGGAGCGCGGACAGGACAAACGGGATTAAGAAAGGCGCCACGCTTTAG	1790
Db	28298	AGATCAGGAGGAGCGCGGACAGGATTAAGAAACGGGATTAAGAAAGGCGCCACGCTTTAG	28240
Qy	1791	TCATGACCTTCAGGCAAG-TGGACTTTTGGAGGCTCTGGAAGGAGGAAAGCTGGGCAAT	1849
Db	28239	TCATGACCTTCAGGCAAGCAAGCTTCGAGGCTCTGGAAGGAGGAAAGCTGGGCAAT	28180
Qy	1850	TGAATGCTTAATAGGCTTGTTCAGTGGGCTTACAAGGACACTTTAAAAAGATTGT	1909
Db	28179	TGAATGCTTAATAGGCTTGTTCAGTGGGCTTACAAGGACACTTTAAAAAGATTGT	28120
Qy	1910	CAAAGTAGAAGTAAAGC-CCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAA	1968
Db	28119	TCTAATAGAAATAAAGCAACCCCTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAA	28060
Qy	1969	GGCCCACTGCCAGGGGACAAAGGCTTTTGTAGTCAGAGGACCACTAAACAGATGATCCA	2028
Db	28059	GGCCCACTGCCAGGGGACAAAGGCTTTTGTAGTCAGAGGACCACTAAACAGATGATCCA	28000
Qy	2029	GCAGCAGGACTGAGGCTGCTGGGCAAGCGGCTATCCCATGCCATCACCCCTCACAGAGCC	2088
Db	27999	GCAGCAGGACTGAGGCTGCTGGGCAAGCGGCTATCCCATGCCATCACCCCTCACAGAGCC	27940
Qy	2089	CTGGGTATGCTTGACCAATTTAGGGCCAGGAAGGT---TGTCTCTGGACACTGGTGGGT	2145
Db	27939	CCGGGTATGCTTAAACCAATTTAGGGCCAGGAAGGTAACTGTCTCTGGACACTGGGCTGSC	27880
Qy	2146	CTTCTTAGTCTTACTCTTGTGTCGGGCAAACTGTCTCTCCAGATCTGCTCACTATTCTGAG	2205
Db	27879	CTTCTTAGTCTTACTCTCTGCTGCTGGACAACTGT-CTCCAGATCTGCTCACTATCCAAG	27821
Qy	2206	GGGTCCTCNTAAGACGGGACGTCACTAGATACTTTTTCAGCCCACTAAGTTATGAATGG	2265
Db	27820	GG-----TCCTGGGACAGCCACTAGATAC-TTCTCCAGCCACTAAGCTGTG-ACTGG	27770
Qy	2266	GGAGCTTATTTCTTTTACATGCTTTTCTAATATATGCTTGAAGCCCACTACCTTTGTTA	2325
Db	27769	GGAACTTTACTCTTTTACATGCTTTTCTAATATATGCTTGAAGCCCACTACCTTTGTTA	27710
Qy	2326	GGGAGAGACATTTCTAGCAAAAGCAGGGCCCAATATACACCTGAACTAGAGAGGAAACA	2385
Db	27709	GGGAAAGACATTTCTAACAAGCAGGACCAATATACACCTGAACTAGAGAGGAAACA	27650
Qy	2386	CCGTTTGTGTGTCCTGCTTGGAGAGGAAATTAATCTCTGAAAGTCTGGGCAACAGAGG	2445
Db	27649	CCGTTTGTGTGT-CCCTGCTTGAAGAGGAAATTAATCTCTGAAAGTCTGGGCAACAGAGG	27591
Qy	2446	ACAATATGGACGAGCAAAAGAAATGCCGCTCTCTCAAGTTAACTAAGAGATTCACATT	2505
Db	27590	ACAATATGGACGAG-CAAAATATGCCGCTCTCTCTAGTTAACTAAGAGATTCACACT	27532
Qy	2506	CTTTTCCCTACCAAGGACGTACCCCTCAGACCCCAAGGCCCAACAGGATTCACAAAGA	2565
Db	27531	CTTTTCCCTACCAAGGACGTACCCCTTAGACCCGAGGCCCAACAGGATTCACAAAGA	27472
Qy	2566	TTGTTAAGGACTTAAAGGCCCAAGGCTTAGTAAACCATATGATCTCTCTGCTGAGTAATT	2625
Db	27471	TTGTTAAGGACTTAAAGGCCCTACTCTCTAGTAAACCATATGATCTCTCTGCTGAGTAATT	27412
Qy	2626	CCGTAGTGGATTGAGAGGACAGAAACCCAGTGGACAGTGGAGGGTTAGTCAAGATCT	2685
Db	27411	C-----AATTTTAGGAGTAAAGAGCCCAATAGACAGTGA-GGTTAGTCAAGATGT	27360
Qy	2686	CAGGATTATCAATGGAGGCGTTGTCTTTTATACCCAGCTGTACCTAGCCCTTATATCTG	2745
Db	27359	CAGGATTATCAGT-GAGGCTGTCTTCTCTATACCCAGCTGTACCTTACCTTCTACTC	27301
Qy	2746	TGMYTTCCAAATACAGAGGAGCAGAGTGGTTTACASTCTCTGGACCTTMAAGATGCCIT	2805
Db	27300	TGCTTTCCCAATACAGAGGAGCAGAGTGGTTTACAGTCTCTGGACCTTMAAGATGCCIT	27241
Qy	2806	TCTTCTGATCTCCCTGTATATCTGATCTCAATCTTGTGTGCTTTTGAAGATCTTCAA	2865
Db	27240	TTTCTGATCTCCCTGTATATCTGATCTCAATCTTGTGTGCTTTTGAAGATCTTCAA	27181
Qy	2866	ACCCATCATCTCACTCAGCTGATCTTTTATACCCCAAGGTTTCAAGGATAGTCCCATC	2925
Db	27180	ACCCATCATCTCACTCAGCTGATCTTTTATACCCCAAGGTTTCAAGGATAGTCCCATC	27121
Qy	2926	TATTTGGCCAGGCAATTTAGCCCAAGCTTGAGYCATYMTATACCTTGGACACTCTTTGTCC	2985
Db	27120	TATTTGGCCAGGCAATTTAGCTCAAGACTTGAGACAGTTCATACCTGAAACAGTCTTTGTCC	27061
Qy	2986	TTTCRTAGTGTGATGATTTACTTTTGTGCTTCTTGTGCTTGTGCTTGTGCTTGTGCTT	3044
Db	27060	TTTGGTACATGATAATCTACTTCTAGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTT	27001
Qy	3045	ACCCAGGCTCTTMAATTTCTCCGYACCTGTGGCTCAWGGTTTCCAAASARAGCT	3104
Db	27000	ACCCAGGCTCTTMAATTTCTCCGYACCTGTGGCTCAWGGTTTCCAAASARAGCT	26941
Qy	3105	CARCTCTGTCTACAGCAGGTTAAATTAATTTAGGCTTAATTAATTTAGGCTTAAATTAAT	3164
Db	26940	CAGCTCTGTCTACAGCAGGTTAAATTAATTTAGGCTTAAATTAATTTAGGCTTAAATTAAT	26881
Qy	3165	CTCAGTGGAGAAATATCAGGCTTATCTGGCTTATCTCTATCTCTATCTCTCTCTCTCT	3224
Db	26880	CTCAGTGGAGAAATATCAGGCTTATCTGGCTTATCTCTATCTCTCTCTCTCTCTCTCT	26821
Qy	3225	CTAAGGRTTCTTGGGRTTAAGGTTCTGCGGAAATGATTTCCCAAGTGTGGCA	3284
Db	26820	ATAAAG-GTTCTTGGGATAACAGGCTTCTGCCAAATATGATTT-CCCAAGTGTGGCA	26763

```
QY 3285 AATGCCAGGYCATTAWATACASTAAATTAAGGAACTCAGAAAGCAATACCCATTART 3344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26762 AATGCCAGGCCATTATATACHTAATTAAGNAGCTCAGAAAGCAATACCCATTAGT 26703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3345 AAGATGGAYMGTGAAGYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAA 3404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26702 AAATGGACTCTCTAAGCAAAAGCAGCTTTCCAGGCCCTTAAAGAGGCCCTT-AAACCCAA 26644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3405 GYCCAGTGTTHAAGYTTGCCAACRGGGCGAAGACTTTTSTTATATYRTACAGAAAAAAC 3464
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26643 GACTCAGTGTTHAAGCTTACCAACGGGGCAAGACTTTTCTCATATGTGCAC---AAAAAAC 26587
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3465 AGRAAYAGCTCTRGAGTCTTACACAGRTCCRAGGGYAGCTTGGCAACCVRTGGCRYA 3524
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26586 AGAAATAGCTCCAGGAGTCTTACACAGTTCNAGGGACCGCTTTCGCAACCATGGCATA 26527
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3525 CCTGASTAAGGAAAYTGTAGTGTAGTGGCAAGGGTTGRCVTCATTTTAYGGGTAGTGGT 3584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26526 CCTAAGTAAGAAATTAATGTAGTGGCAAGGGTTGGCATCATTTTACGGGTAGTGGC 26467
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3585 GGCAGTAGCAGHYKTAGTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGT 3644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26466 GGCAGTAGCAGTCTTAGTATGTAAAGCAGTTTAAATAATACAGGGAAGAGATCTTACTGT 26407
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3645 GTGGACATCTCATGAKGTGAAYRGCATCTCACTGCTAAAGGAGACTTGTGGCTGCAGA 3704
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26406 GTGGACATCTCAATGATGTAAACAGCATCTCACTGCTAAAGGAGACTTGTGGCTGCAGA 26347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3705 CAACYGTTTACTTAATRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCAC 3764
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26346 CAATCGTTTGTCTAATATCAGCTCTATTAAATAAGGCCAGTGTGCAACTGCATAC 26287
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3765 TTGTGCACTCTTAACCCAGYCNCAITTTCTCAGACAATGAAGAAAAAGATARAAYATA 3824
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26286 TTGCGCACTCTTAACCCCGCACATTTCTTCAAAACAATGA-AAAAGATAAAACATAA 26229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3825 CTGTCAACAARTAAATTTCAAACTATGCACTCGAGGGACCTTGTAGAGTTCCTTT 3884
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26228 CTGTCAACAGGTAATTTTCAAACTATACCACTCGA-GGGACCTTCTAGAGTTCCTTT 26170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3885 GACTGATCCYGACCTTCAACTTGTATGATGGAAGTTTCTTTGTAGAAAAAGGACTTC 3944
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26169 GACTAATCCCTAAC-TCACTTGTATATTAATGAAAATTCCTTTGTAAAAAAGGACTTC 26111
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3945 GAAAAGYGGGTATGCAGTGGTCAGTGATGATGGAATATTTGAAAGTAAATCCCTCACTC 4004
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26110 -AAAAAGCGAGGTATGCAGTGGTCAGTAAATGGAATATCTTGAAGTAAATCCCTCACTC 26052
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4005 CAGGAACTAGTCTYAGCTRGCAAGTAATAGCCVTCATYTKGGGCACTAGAAATTAGGAG 4064
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26051 CAGGAACTAGTCTCGGCTGGCAGAACTAATAGCCCTCACTCGGGCACTAGAAATTAANA 25992
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4065 AAGRAAAAAAGGYAATATATATACAGACTCTTARTATGCTYACCTAGTCNTCCATGCC 4124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25991 AAGGNAAGGTAATATACATACAGACTCAAGTATTTTACCTAGTCTCTACGCC 25932
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4125 ATMRGCAATATGSPAPAGAAAGGAATTCCTAACTTCYAGRGAAACCTCTATCAMACATC 4184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25931 ACACAGCAATATGGAAGAAAAAGGAATTCCTAACTTCTGAGGGAACACCTATCAAAATC 25872
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4185 AGGAGCCATTAGGABATTTATYTGCGWGTACAGAAACCTAPAGAGGTGGMAGTCTTAC 4244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25871 AGGAGCCATTAGGAGATTTATTTGGCTGTACAGAAACCTAAGAGGTTGGCAGTCTTAC 25812
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4245 ACTGCVGGGTATCATCANAAAGGAAGAAAGGGAATASAGRAAYTCCCAAGCAKATA 4304
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25811 ACTGCGGGGTATCATGAAAGAAAGAAAGAAATAAGAACTCCCAAGCGGTA 25752
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4305 TTGAAGCMAAAGAGCTGCAAGCAGGACCTCCATTAGAAATGCTTATTAACCTTCCCT 4364
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25751 TTAAGCTAAAAAGAGCAAGGAGAACTCTCCATTAAAAATGCTTATAGAGGAGCCCC 25692
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4365 TAGTATAGGTAATCCCTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAAT 4424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 25691 TAGTATGGGTAAATCCCTCTTGGGNAACCAAGTCG--TACTCAGCAGGAGAAATAGAAC 25634
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4425 GGGGAACCTCAC---GAGGCAAGTTTTTCCTCCTCGGACGGTGTAGCCACTGAAAGAGGA 4481
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25633 GGGGAACCTCAAGAGGAGTAGTTCCTCCTCCTCAGAAATGGCTAGCCACCGAAAAAGGAA 25574
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4482 AAATACTTTTCCCTGCAACTATCCAATGGAATTTACTTTAAACCTTTCATCAAACTTTTC 4541
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25573 AAATACTTTTCCCTGCAACTAACAANTAGAANTTACTTTAAACCTTTCACCAACTTTTC 25514
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4542 ACTTAGGCATCGATAGCACCCCATCARATGGCCAAATCAATTTTACTGGACCCAGGCTTTT 4601
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25513 ACTTAGGCATTAATAGCACCCCATCAGATGGCCAAATCATTTTACTGGACCCAGCTTTT 25454
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4602 TCAAAACTATCAAGCARATATKAGGGCTGTGAATGTGTCARARAATAATATCCCTGTC 4661
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25453 TCAAAACTTAAACAGCAGATAGTCAGGGCTGTGAAGTGTGTCAAAAAAATAACCCCTGTC 25394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4662 CTYATCGCAAGCTCCTTCA 4681
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25393 ACTGAGGCCATACATTTCA 25374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
US-09-719-554-1
; Sequence 1, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH H
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-1

Query Match 34.1%; Score 2550.4; DB 3; Length 2599;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2552; Conservative 39; Mismatches 5; Indels 2; Gaps 2;

QY 4655 CCCCTGCTCTYATCGCCAAAGTCTCTTCAGAGARAAACAAARAAACAGGCCATTTACCTGPARAA 4714
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 CCCCTGCTCTAATCGCCAAAGTCTCTTCAGGAGAAACAAAGAACAGGCCATTTACCTGGAGAA 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4715 RACTGGCACTGATTTTACCACAGCCCAACCTTCAGGATTTTCAGTATCTTACTAGTCT 4774
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 GACTGGCACTGATTTTACCCCAAGCCCAACCTTCAGGATTTTCAGTATCTTACTAGTCT 123
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4775 GGGTARATATCTTTCACGGGTTGGGCGARAGGCCCTTCCCTGTAGGACAGAAAAAGGCCCAAG 4834
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 GGGTAGATATCTTTCACGGGTTGGGCGAGGCCCTTCCCTGTAGGACAGAAAAAGGCCCAAG 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4835 AGGTAAATAAGGCACATAGTTTCATGAAATAATTTCCAGATTCGGACTTCCCGAGGCTTAC 4894
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 AGGTAAATAAGGCACATAGTTTCATGAAATAATTTCCAGATTCGGACTTCCCGAGGCTTAC 243
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4895 AGAGTGACAATAGCCCTGCTTCCAGGCCACAGTAACCCAGGAGTATCCAGGCCCTTAG 4954
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 AGAGTGACAATAGCCCTGCTTCCAGGCCACAGTAACCCAGGAGTATCCAGGCCCTTAG 303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 4955 GTATACGATATCACTTACACTCGCGCTGAAAGGCCACAGTCTCTCAGGGAAGGTCGAGAAA 5014
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



Db 304 GTATACGATATCACTTACCTGCGCTGAAGGCCACAGTCTCAGGGAAGGTGAGAAAA 363  
Qy 5015 TGAATGAAYACTCAAGAGCATCTAAAAAGAAAGAAACCCAGGAACCCACCTCACATGGC 5074  
Db 364 TGAATGAACACTCAAGAGCATCTAAAAAGAAAGAAACCCAGGAACCCACCTCACATGGC 423  
Qy 5075 CTGTCCTGTTGCGCTATAGCCCTTAAAAAGAAATCTGCAACTTTCCCAAAAAAGCAGGACTTA 5134  
Db 424 CTGCTCTGTTGCGCTATAGCCCTTAAAAAGAAATCTGCAACTTTCCCAAAAAAGCAGGACTTA 483  
Qy 5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCAATAACCAATGACCTTTGTGCTTGACCCAA 5194  
Db 484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCAATAACCAATGACCTTTGTGCTTGACCCAA 543  
Qy 5195 GACAGCCAACTTAGTTCAGACATCACTCTCTTAGCCAAATATCAACAAGTTCTTAAAC 5254  
Db 544 GACAGCCAACTTAGTTCAGACATCACTCTCTTAGCCAAATATCAACAAGTTCTTAAAC 603  
Qy 5255 ATTACAAGGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTGAGCATGGT 5314  
Db 604 ATTACAAGGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTGAGCATGGT 663  
Qy 5315 ATTAGTCAAGTCCCTTCYCTTAATCCCATCCCTAGATACATCTCTGGGAAGACCCCTA 5374  
Db 664 ATTAGTCAAGTCCCTTCCTCTTAATCCCATCCCTAGATACATCTCTGGGAAGACCCCTA 723  
Qy 5375 CCAGTCATTTTATYATACCCCAACTGCGGTAAAGTGGCTGGAGTGAGTCTTGGATACA 5434  
Db 724 CCAGTCATTTTATYATACCCCAACTGCGGTAAAGTGGCTGGAGTGAGTCTTGGATACA 783  
Qy 5435 TCACACTGAGTCAAACTCTGAGATCTGCGCAAGGAACTGAAATCCAGGAGACAACGC 5494  
Db 784 TCACACTGAGTCAAACTCTGAGATCTGCGCAAGGAACTGAAATCCAGGAGACAACGC 843  
Qy 5495 TAGCTATTCTGTGAACCTCTAGAGGATTTGCGCTGCTCTTCAAAACAAACACAGGAGG 5554  
Db 844 TAGCTATTCTGTGAACCTCTAGAGGATTTGCGCTGCTCTTCAAAACAAACACAGGAGG 903  
Qy 5555 AAGTAACTAAATCAATAATCCCATGAGGCTTCCCTTATCATATTTTTCTCTKTAAGT 5614  
Db 904 AAGTAACTAAATCAATAAT - CCCCATGCCCCCTCCCTTATCATATTTTTCTCTTACATG 962  
Qy 5615 TTSTTTTACCTCTTCACTCTCACTGCAACCCCTCCATGCGCTGTATGACAGTAGCT 5674  
Db 963 TTCTTTTACCTCTTCACTCTCACTGCAACCCCTCCATGCGCTGTATGACAGTAGCT 1022  
Qy 5675 CCCTTACCMAGAGTTTCTATGAGAAATGAGAGTCCCGAAATATTGATGCCCCCATCGT 5734  
Db 1023 CCCTTACCMAGAGTTTCTATGAGAAATGAGAGTCCCGAAATATTGATGCCCCCATCGT 1082  
Qy 5735 ATAGGAGTCTTTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCGCAACT 5794  
Db 1083 ATAGGAGTCTTTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCGCAACT 1142  
Qy 5795 GCTATCACTCTGCACTCTTTGATGATGATGCAAAATCAATATTGGAAGGAAATGA 5854  
Db 1143 GCTATCACTCTGCACTCTTTGATGATGATGCAAAATCAATATTGGAAGGAAATGA 1202  
Qy 5855 TTAATCTAGTCTGCTGAGGAGTCTGAGTCACTGCTGCTGGAATCTTCACTTCAACCCAA 5914  
Db 1203 TTAATCTAGTCTGCTGAGGAGTCTGAGTCACTGCTGCTGGAATCTTCACTTCAACCCAA 1262  
Qy 5915 CTGGTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGGAAAAACATGTAAAAAGAG 5974  
Db 1263 CTGGTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGGAAAAACATGTAAAAAGAG 1322  
Qy 5975 TAATCTCCCAACTCAACCGGGTACATGGCACTCTTAGCCCTTACAAAGGACTAGATCTCT 6034  
Db 1323 TAATCTCCCAACTCAACCGGGTACATGGCACTCTTAGCCCTTACAAAGGACTAGATCTCT 1382  
Qy 6035 CAAAACTACATGAACCCCTCCGTACCCATCTGCTGCTGTAAGCCTATTTAATACACCC 6094  
Db 1383 CAAAACTACATGAACCCCTCCGTACCCATCTGCTGCTGTAAGCCTATTTAATACACCC 1442

Qy 6095 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTGATATGCTCCCCC 6154  
Db 1443 TCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTGATATGCTCCCCC 1502  
Qy 6155 TGAACCTTCAAGCCCATATGTTTCAATCCCTGACTGAAACAACTTCAAGCAAG 6214  
Db 1503 TGAACCTTCAAGCCCATATGTTTCAATCCCTGACTGAAACAACTTCAAGCAAG 1562  
Qy 6215 AAATAAACACCACTTCCGTTTATAGTAGGACCTCTTGTTCCTCAATSTGAAATAACCCATA 6274  
Db 1563 AAATAAACACCACTTCCGTTTATAGTAGGACCTCTTGTTCCTCAATCTGAAATAACCCATA 1622  
Qy 6275 CCTCAAAACCTCACTCTGTGTAAATAATTTAGCAATACTACATACAAACCACTTCCCAATGCA 6334  
Db 1623 CCTCAAAACCTCACTCTGTGTAAATAATTTAGCAATACTACATACAAACCACTTCCCAATGCA 1682  
Qy 6335 TCAGGTGGGTAACTCTCCACACAATAAGTCTGCGCTACCCCTCAGGAAATATTTTTGTCT 6394  
Db 1683 TCAGGTGGGTAACTCTCTCCACACAATAAGTCTGCGCTACCCCTCAGGAAATATTTTTGTCT 1742  
Qy 6395 GTGCTACCTCAGCCTATCGTTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAT 6454  
Db 1743 GTGCTACCTCAGCCTATCGTTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAT 1802  
Qy 6455 TCTTAGTGCCCCYATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCTCATATCTA 6514  
Db 1803 TCTTAGTGCCCCYATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCTCATATCTA 1862  
Qy 6515 AGCCCCGCAACAAAAGAGTACCCATTTCTTCTTTTGTATAGGAGAGGAGTGTAGGTG 6574  
Db 1863 AGCCCCGCAACAAAAGAGTACCCATTTCTTCTTTTGTATAGGAGAGGAGTGTAGGTG 1922  
Qy 6575 CACTAGTACTGTCATTTGGCGTATCAACACTCTACTCAGTTCTACTCAAACTATCTC 6634  
Db 1923 CACTAGTACTGTCATTTGGCGTATCAACACTCTACTCAGTTCTACTCAAACTATCTC 1982  
Qy 6635 AAGAACTAAATGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTCCAAAGATCAAC 6694  
Db 1983 AAGAACTAAATGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTCCAAAGATCAAC 2042  
Qy 6695 TTAACCTCCCTAGCAGCAGTAGTCTTTCRAAATCGAAGAGCTTTAGACTGCTAACCGCTG 6754  
Db 2043 TTAACCTCCCTAGCAGCAGTAGTCTTTCRAAATCGAAGAGCTTTAGACTGCTAACCGCTG 2102  
Qy 6755 ARAGAGGGGAACTCTGTTTATTTTAGGGGAAAGTCTGTTATTTAGTTAACTCAATCCG 6814  
Db 2103 AAGAGGGGAACTCTGTTTATTTTAGGGGAAAGTCTGTTATTTAGTTAACTCAATCCG 2162  
Qy 6815 GAATCTGCTCACTGAGAAAGTTTAAAGAAATTCAGATCGAAATCAAACTGAGCAGAGAGC 6874  
Db 2163 GAATCTGCTCACTGAGAAAGTTTAAAGAAATTCAGATCGAAATCAAACTGAGCAGAGAGC 2222  
Qy 6875 TTCGAAACACTGGAACCTCGGGCTCTCAGCCATGAGTGGCTGGAATCTCCCTTCT 6934  
Db 2223 TTCGAAACACTGGAACCTCGGGCTCTCAGCCATGAGTGGCTGGAATCTCCCTTCT 2282  
Qy 6935 TAGGACCTCAGCAGCTAATAATTTGCTACTCTCTTTGGACCTGATCTTTTACCTCC 6994  
Db 2283 TAGGACCTCAGCAGCTAATAATTTGCTACTCTCTCTTTGGACCTGATCTTTTACCTCC 2342  
Qy 6995 TTGTTAACTTTGCTCTTCCAGAACTGTAAGTGTAAACTTACAAATGGAGCCCAAGATGC 7054  
Db 2343 TTGTTAACTTTGCTCTTCCAGAACTGTAAGTGTAAACTTACAAATGGAGCCCAAGATGC 2402  
Qy 7055 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGAACCGGCTGTAGCCCAAGATCTGATG 7114  
Db 2403 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGAACCGGCTGTAGCCCAAGATCTGATG 2462  
Qy 7115 TTAATGACATCAAGGACCCCTCTGAGGAAATCTCAGCTGCAAACTCTTACTAGCC 7174  
Db 2463 TTAATGACATCAAGGACCCCTCTGAGGAAATCTCAGCTGCAAACTCTTACTAGCC 2522



```

QY 7175 CCAATTCAGCAGGACGTTAGAGCGTSGTGGCCCACTCCCAACAGCACTTAGGT 7234
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2523 CCAATTCAGCAGGACGTTAGAGCGGTC-TCGGCCAACTCCCAACAGCACTTAGGT 2581
QY 7235 TTTCCTGTTGAGATGGG 7252
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2582 TTTCCTGTTGAGATGGG 2599

RESULT 10
US-09-719-554-4
; Sequence 4, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2784
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-4

Query Match 30.5%; Score 2282.4; DB 3; Length 2784;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 2523; Conservative 39; Mismatches 180; Indels 66; Gaps 9;

QY 4674 CTCCTTCAGGAAACAAAPAAACAGGCCATTACCTGGRAPAAABACTGGCAACT-GATTTTA 4732
Db 1 CTCCTTCAGGAAACAAAPAAACAGGCCACTACCAAGAGAAAGACTGGCAACTAGATTTTA 60

QY 4733 CCCAAGGCCAAACCTCAGGAGTTTCAGTATCTACTAGTCTGGGTARATCTTTACCGG 4792
Db 61 CCATATGCCAAATCTCAGGATTCAGTATCTACTAGTTGGGTAGATCTTTCACGT 120

QY 4793 GTTGGCAGAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAGGCACTAG 4852
Db 121 GTTGGCAGAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAA- ----CG 174

QY 4853 TTCATGAATAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTG 4912
Db 175 TTCATGAATAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTG 234

QY 4913 CTTTCCAGGCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTAC 4972
Db 235 CTTTCAAGGCTACAGTAACCCAGGAGTATCCAGGTTAGGTATACATATCACTAC 294

QY 4973 ACTGGCCCTGAAGGCCACAGTCTCAGGGAAGTCCGAGAAAATGAATGAAYACTCAAG 5032
Db 295 ACTGGCCCTGGAGGCCACAGTCTCAGGAAAGGTGGAGAAAATGAACAAACACTCAAA 354

QY 5033 GACATCTAAAGAAAGCAACCCAGGAACCCACTCAGATGGCTGTCTGTTGCCCTATAG 5092
Db 355 GACATCTAAAGAAAGCAATCCAGGAACCCACTCAGATGGCTGTCTGTTGCCCTATAG 414

QY 5093 CCTTAAAGAAATCTGCAACTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGT 5152
Db 415 CTTTACTAAGATCCGAAATCTCCCAAAAGCAGGACTTAGTCCATACAAATGCTGT 474

QY 5153 ATGGAAGGCCCTTCATPACCAATGACCTTGTGCTGACCCA-AGACAGCCAACTTAGTTG 5211
Db 475 ATGGAAGGCCCTTCCTTAAACCAATGAATTTGGGCTTGACCGAGAGACAGCAACTTAGTTG 534

QY 5212 CAGACATCACTCCTTAGCCAAATATCAACAGTTCTTAAACATTTACAAGNACCTATC 5271

```

```

Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
535 CAGACATCATCTCTTAGCCAAATATCAACAGGTTCTTAAACATTTACAGGGAGCGTGT 594
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5272 CTTGAGAAGAGGAAAAGAACTATTCCACCWGTGACATGGTATTAGTCAAGTCCCTTC 5331
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
595 CCAAGAAGAGGAAAAGAACTATTCCACCCTGGTGAATGTTAGTCAAGTCCCTTC 654
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5332 YCTCTAATTTCCCATCCCTAGATACATCTCGGAAGAGCCCTACCAGCTCATTTTATTA 5391
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
655 CTTCTAATTTCCCATCCCTAGATACATCTCGGAAGGAACTACCAGCCATTTTATCTTA 714
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5392 CCCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTCAAGTCAAT 5451
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
715 CCTTAAACGGCAGTTAAAGTGGCTGGAGCGAGTCTTTGGATACATCACACTCAAGTCAAT 774
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5452 CTTGATATCTGCCAAAGAACTGAAATTCAGAGGAGCAACGCTAGCTATTCTCTGTGAAC 5511
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
775 CTTGATATCTGCCAAAGAACTGAAATTCAGAGGAGCAACGCTAGCTATTCTCTGTGAAC 834
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5512 CTTGATATCTGCCAAAGAACTGAAATTCAGAGGAGCAACGCTAGCTATTCTCTGTGAAC 5571
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
835 CTTGATATCTGCCAAAGAACTGAAATTCAGAGGAGCAACGCTAGCTATTCTCTGTGAAC 894
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5572 AAATCCCCCATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTC 5631
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
895 AAATCCCC--TGGCCCTCCCTTATCATATTTTCTCTTACTGTTCTCTTACCCCTTTC 952
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5632 ACTCTCACTGCAACCCCTCCATGCGCTGTA-----TGACCAGTAG 5672
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
953 ACTCTCACTGCAACCCCTCCATGCGCTGCAACCCCTCCATGCGCTGCAACCCCTCCATGCGCTGCA 1012
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5673 CTCCTCTYACWAGAGTTTCTATGGAGATCGAGGTCGCGGAATATTTGATGCCCATC 5732
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1013 CTCCTCTYACWAGAGTTTCTATGGAGATCGAGGTCGCGGAATATTTGATGCCCATC 1072
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5733 GTATAGGAGTCTTTTAAAGGAAACCCCACTTCTACTGCCACACCCATATGCCCGCAA 5792
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1073 GTATAGGAGTCTTTTAAAGGAAACCCCACTTCTACTGCCACACCCATATGCCCGCAA 1132
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5793 GTGCTATCACTCTGCACTCTTTGATGATGCAATTAATCTCATTTATGGACAGGAAAAT 5852
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1133 GTGCTATCACTCTGCACTCTTTGATGATGCAATTAATCTCATTTATGGACAGGAAAAT 1192
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5853 GATTAATCTTATGCTGCTGGAGGACTTGGAGTCACTGCTGCTGGAGTCTTCTTACCCCA 5912
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1193 GATTAATCTTATGCTGCTGGAGGACTTGGAGTCACTGCTGCTGGAGTCTTCTTACCCCA 1239
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5913 AACTGCTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTTAAAGA 5972
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1240 TACCAGTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTTAAAGA 1299
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5973 AGTAATCTCCCAACTCAACGSGGTACATGGGACCTCTAGCCCCCTCAAGAGGACTAGTCT 6032
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1300 AGTAATCTCCCAACTCAACGSGGTACATGGGACCTCTAGCCCCCTCAAGAGGACTAGTCT 1359
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6033 CTCAAAACCTACATGAAACCCCTCCGTCACCATCTCGCTGGTAAAGCTATTTAAATACC 6092
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1360 CTCAAAACCTACATGAAACCCCTCCGTCACCATCTCGCTGGTAAAGCTATTTAAATACC 1419
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6093 CCTCACTGGGCTCCATGAGGCTCTCGGCCCAAAACCTCTACTAATCTTGGATATGCTCTCC 6152
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1420 CCTGATCTGGGCTCCATGAGGCTCTCGGCCCAAAACCTCTACTACTTGTGATGCTCTCC 1479
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6153 CCTGAATCTCARGCCATATGTTTCAATCCCTGTAACCTGAACAAATGGAAACAACTTCAGCAC 6212
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1480 CCTGACCTTTAGGCCATACATTTCAATCCCTATACCTGAAACAACTTCAGCAC 1539
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6213 AGAATAAACACCACTTCGTTTATAGGACCTCTTGTTCATSTCCATSTCGAATAACCCA 6272
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1540 AGAATAAACACCACTTCGTTTATAGGACCTCTTGTTCATSTCCATSTCGAATAACCCA 1596
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6273 TACCTCAAAACCTCACTGTTGTAATAATTTAGCAATACATACACCAACCACTCCCAATG 6332

```



Db	96607	AAATAGACCTAGGTAATTTCTCAGATACCCCTTAATGGCTATATTTGATGTGTTTTACAAGGTT	96548
Qy	1169	TAGGCAATTTCTTGATCTGACATCGAGAGATATATGTCTCTCTAAATCAGACACTA	1228
Db	96547	TAGGCAATTTCTTGATCTGATATGGAGATATA-ATGTTACTCTCTAAATCAGACACTA	96489
Qy	1229	ACCCCAATGAGAAAGTGGCCACCAATTAATCGAGCTGAGGTGTTGGCGATCTCTGGTAT	1288
Db	96488	ACCCCAATGAGAAAGTGTGCGGTAACTGAGCTTGAAGTTTGGCGATCTCTGGTAT	96429
Qy	1289	CTCAGTCAGGTCAATGATGATGATGACAAACAGAAAGAAAGANAATGATTCGCCACAGGC	1348
Db	96428	CTCAGTCAGGTCAATG- --ATAGGTCGACACAGAGAAAGAGACGATTCGCCACAGGC	96372
Qy	1349	CAGCARGCAGTTCOCAGTCTASACCTCATTTGGGGACACAGAAATCAGTAACATGGGAGA	1408
Db	96371	CAGCAGGCAGTTCOCAGTGTAGACCTCTACTGGG- --ACACAGAAATCAGAAAT- --GGAGA	96316
Qy	1409	TGCTGCTGCAGACATTTGCTAACTTGTGCTASAGGACTTAAGGAAACTTASGAGAA	1468
Db	96315	TTGGTGGCGGACACATTTGCTAACTTGTGCTAGAGGACTTAAGGAAACTTASGAGAA- A	96257
Qy	1469	ARTCTAYGAATTTACTCAATGATGTCCACCATAAACAGGGGAGGAAAGAAATCTCTACT	1528
Db	96256	AGCCTGTGATTTTCAATGATGTCCATATNAACACA- --GGNAAGGAGAAATCTCTACT	96198
Qy	1529	GCCTTCTGGAGAGACTAAGGGAGGACTTGAAGAGCGTGCGCTCT- --CTGTCACTGACT	1586
Db	96197	GCCTTCTGGAGTACTAAGCAGGCAATTTAGGAAAGTACCTCTCTCTGCTCACTGACT	96138
Qy	1587	CTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTTTATCACTCAGTCAGTCGAGACATTA	1646
Db	96137	CTACTGAAGGCCAACTAATCTTTAAAGGATAAGTTTATCACTCAGTCAGTCAGACATTA	96078
Qy	1647	-GAAAAAATCTCAAAAGTCTGCGTAGGCCGCGAGCAAACTTTAGAAACCTTATTGAAT	1705
Db	96077	GGAAAAAATCTCAAAAGTCTGCGTTAGGCCGCGAGCAAACTTTAGAAACCTTATTGAAT	96018
Qy	1706	TGGCAACTTCGTTTTTTTATATAGAGATCAGGAGGAGCAGCGGAAACAGCAAAACGGG	1765
Db	96017	TGGCAACTTCGTTTTTTTATATAGAGATCAGGAGGAGCAGCGGAAATGGGACAAATGGG	95958
Qy	1766	ATTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCAAGTGGACTTTGGAGGCTCT	1825
Db	95957	A-TAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCAAGTGGACTTTGGAGGCTCT	95899
Qy	1826	GGAAAAAGGAAAGCTGGGCAAAATGAATGCGCTTAATAGGCTTGTCTCCAGTGGGCTTA	1885
Db	95898	GGAAAAAGGAAAGCTAGGCAAAATCAATGCGCTTAATAGGCTTGTCTCCAGTGGGCTTA	95839
Qy	1886	CAAGGACACTTTAAAAAGATTGTCCAAGTAGAAGTAAAGCCGCCCTTGTCTCCATGCCCC	1945
Db	95838	CAAGGACACTTTAAAAAGATTGTCCAATAGAAATAAGCCGCCCTTGTCTCCATGCCCC	95779
Qy	1946	TTATTTCAAGGAACTCATGGAAGGCCCTCTGCCCGAGGAGCAAAAGTCTTTTGAATCA	2005
Db	95778	TGCTGTCAAGGAACTCATGTAAGGCCCTCTGCCCGAGGAGCAGTAGGCTCTCTGAGTCA	95719
Qy	2006	GAGGCCACTAACAGATGATCAGCAGCAGGACTGAGGCTGCTGGGCAAGGCCATCC	2065
Db	95718	GAGGCCACTAACAGATGATCAGCAGCAGGACTGAGGCTGCTGGGCAAGGCCATCC	95659
Qy	2066	CATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGCCGAGGAAGGT- --	2122
Db	95658	CATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGCCGAGGAGTAAAC	95599
Qy	2123	TGCTCTCTGGACACTGGTGGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2182
Db	95598	TGCTCTCTGGACACTGGTGGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	95539
Qy	2183	TCCAGATCTGTCACTATTCTGAGGGGCTCCNTAAGACGGGAGTCACTAGATCTTTTTC	2242
Db	95538	TCCAGATCTGTCACTATCC- --AAGGGGCTCTAGGACAGCCAGTCACTAGATAC- --TTCTC	95483
Qy	2243	CCAGCCACTAAGTTATGAACCTGGAGGCTTTATTTCTTTTTCACATGCTTTTCTAAATATGC	2302
Db	95482	CCAGCCACTAAGTTATG- --ACTGGGAACTTCACTCTTTTTCACATGCTTTTCTAAATATGC	95424
Qy	2303	TTGAAAGCCCACTACCTTTGTTAGGAGAGACATTTAGCAAAAGCAGGGGCACTATATAC	2362
Db	95423	CTGAAAGCCCACTCTCTTTGTTAGGAGAGACATTTAGCAAAAGCAGGGGCACTATATAC	95364
Qy	2363	ACCTGAAACATAGGAGAAACACCCGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2422
Db	95363	ACCTGAAACATAGGA- --GAACACCCGTTTGTGTT- --CCCTGCTTGAAGAGAAATTAAT	95309
Qy	2423	CTGAAAGTCTGGCAACAGAAAGACATATGACGAGCAAAAGATGCCCTGCTTCTTCA	2482
Db	95308	CTTGAAGTCTGGCAACAGAAAGACATATGACGAG- --CAAAGATGCCCTGCTTCTTCA	95250
Qy	2483	AGTTTAAACTAAGGATTCACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2542
Db	95249	AGTTTAAACTAAGGATTCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	95190
Qy	2543	GGCCCAACAGGATTTCAAAGATTTGTTAAAGACTTTAAAGCCCAAGGCTTAGTAAAAAC	2602
Db	95189	GGCTCAACAGGACTTCCAAAAGA- --TTAAGGACCTAAAGGCCCAAGGCTTAGTAAAGC	95133
Qy	2603	ATGCATAACTCCCTGACGATTAATTCGTTAGTGGATTGAGGAGCAGCAAAACCCAGTGCAC	2662
Db	95132	ATGCAATAAGCCCTTACATTAATCC- --AACTTTAGGAGTACAGAAACCCAGTGCAC	95080
Qy	2663	AGTGAGGCTTGTGCAAGATCTCAGGATTAATCAATGAGGCGGTTGCTCTTTTATACCC	2722
Db	95079	AGTGA- --GTTTGTGCAAGATCTCAGGATTAATCAAT- --GAGTCACTGTCTCTTATACCT	95022
Qy	2723	AGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATACACAGAGGAGCAGAGTGGTTTAC	2782
Db	95021	AGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATACACAGAGGAGCAGAGTGGTTTAC	94962
Qy	2783	ASTCTGACCTTGAAGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2842
Db	94961	AGCTGACCTTGAAGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	94902
Qy	2843	GTTTGCCTTTGAAGATCTTCAAAACCCACATCTCAACTCAGCTGAGCTTCTTATACCCCA	2902
Db	94901	ATTGCTCTTTGAAGATCTTCAAAACCCACATCTCAACTCAGCTGAGCTTCTTATACCCCA	94842
Qy	2903	AGGTTTCAAGGATAGYCCCTCATCTATTGTCGAGGACATTTAGCCCAAGACTTCAGYCATY	2962
Db	94841	AGGTTTCAAGGATAGYCCCTCATCTATTGTCGAGGACATTTAGCCCAAGACTTCAGYCATY	94782
Qy	2963	MTACATCTGACACTCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3020
Db	94781	CTCATCTGACACTCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	94722
Qy	3021	TCAGAAACCTTGTGCTCATCAAGCAGCCCAAGCTTAAATTTCTTCTGCTGCTGCTGCTGCT	3080
Db	94721	TCAGAAACCTTGTGCTCATCAAGCAGCCCAAGCTTAAATTTCTTCTGCTGCTGCTGCTGCT	94662
Qy	3081	TACAGGTTTCCAAACSAARAGCTCARCTCTGCTCAGCAGGCTTAAATTAATCTAGGCTTA	3140
Db	94661	TACAGGTTTCCAAACSAARAGCTCARCTCTGCTCAGCAGGCTTAAATTAATCTAGGCTTA	94602
Qy	3141	APATATTCAGAGGACCAAGGCCCTCAGTGAAGAAATATCCAGCTTACTTCTGCTTAT	3200
Db	94601	AAATATTCAGAGGACCAAGGCCCTCAGTGAAGAAATATCCAGCTTACTTCTGCTTAT	94556
Qy	3201	CCTCATCTCAAAACCTTAAAGCAACTAAGGCTTCTTGGCTTAAAGCTTCTGCTGCTGCTGCT	3260
Db	94555	CCTCATCTCAAAACCTTAAAGCAACTAAGGCTTCTTGGCTTAAAGCTTCTGCTGCTGCTGCT	94496
Qy	3261	AWATGATTTCCCGAGTGTGCGAAATAGCCAGGCTTAAATTAATTAATTAATTAATTAATTA	3320
Db	94495	ATATGATTT- --CCAGGATCGGCAAAATAGCCAGCACTTATATACACTAATTAATTAAGGAAAC	94437

Qy	3321	TCAGAAAGCCAAATACCCATTTTARTAGATGGAYACTGAAGYMPAAAGTGGCTTTCCAGGC	3380
Db	94436	TCAGAAAGCCAAATACCCATTTTARTAGATGGAYACTGAAGYMPAAAGTGGCTTTCCAGGC	94377
Qy	3381	CCCTAAAGAACGSCCTTAAACCCCAAGYCCCAAGTGTAAAGYTTGGCAACGGGCAAGACTTTT	3440
Db	94376	CCTAAAGNAGGCCCT--AACCAGCCCAAGTGTAAAGYTTGGCAACGGGCAAGACTTTT	94319
Qy	3441	TSTTVATAYRTCA CAGAAAAA CAGAAAYAGCTCTGGAGTCTTACACAGRTCCRAGG	3500
Db	94318	TCCTTACATGTCA CAGAAAAA CAGAAATAGCTCTAGGAGTCTTACACAGGTC-----	94264
Qy	3501	GAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYTCA TGTAGTGCGAAGGGTTG	3560
Db	94263	GATGAGCTTGCAACCCATGGGCATACCTTGAGTAAAGGAAATTTGATGTAGTGCGAAGGGTTG	94204
Qy	3561	RCYTCATTTTAYGGGTAGTGGTGGCAGTAGCAGTAKTKTAGTATCTCAAGCAGTTAAAT	3620
Db	94203	GCCTCATTTTATGGGTAGTGGCGGAGTAGCAGTCTTAGTATCTGAGCAGTTAAAT	94144
Qy	3621	AATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCTACTCACTGC	3680
Db	94143	AATACAAGGAAGAGAT----CTGTGTAGACATCTCATTAATGTGAACGGCATACTCACTGC	94088
Qy	3681	TAAAGGAGACTTGTGGCTGTGCAGACAACGCT	3711
Db	94087	TAAAGGAGACTTGTGGCTGTGCAGACAACCGT	94057

RESULT 12  
 US-09-719-554-66  
 ; Sequence 66, Application US/09719554  
 ; Patent No. 6919438  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALLIEL, Patrick  
 ; APPLICANT: PERIN, Jean-Pierre  
 ; APPLICANT: RIEGER, Francois  
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU  
 ; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS  
 ; FILE REFERENCE: 200936USOPCT  
 ; CURRENT APPLICATION NUMBER: US/09/719,554  
 ; CURRENT FILING DATE: 2001-12-26  
 ; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
 ; PRIOR FILING DATE: 1999-06-25  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 66  
 ; LENGTH: 46340  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-719-554-66

Query Match	29.6%	Score 2216;	DB 3;	Length 46340;
Best Local Similarity	86.5%;	Pred. No. 0;		
Matches 2726;	Conservative	91;	Mismatches 219;	Indels 115; Gaps 32;
Qy	606	GATGGGAAACGTTCCCGCAGACAAACCGCCCTAAAGCAGTATTCTGGARAAATTTGGGA	665	
Db	28745	GATAGGAAAGGTTCCCTCAGGCGAAACACCCCTAAGATGATTTCTGAGAAATTTGGGA	28804	
Qy	666	MCAATTTGACCCCTCAGACACTAAGAAAGAACGACTTATATCTTCTGCAGTCCGCCCTG	725	
Db	28805	CCAATTTGACTCTCAGATGCTAAGAAAAAAGAAC--ATATTCTTCTGCAGTACCCGCTG	28862	
Qy	726	GCA-----CTCCTGAGGGAAGTATATAATTTAA	753	
Db	28863	GCAACGATATACTCTTTAAGGGGGAGAAACCTGGCATCTCTGAGGGAAGCATATAATTTAA	28922	
Qy	754	CACCATCTTACAGCTAGACTCTTTTCTAGAA---AAGGCAAAATGGAGTGAAGTGCATA	810	
Db	28923	CACCACTTACAGCTAGACCTCTTTTGTAGAAAGAGGCAAAATGGTGTGAAGTGTCA	28982	
Qy	811	AGTACAAACTTTCTTTTTCATTAAAGAGCAACTCACAAATATGTAAAAAGTGTGATTATG	870	

Db	28983	CGTACAAACTTTCTTTTTCATTAAGAGACAACCTCGCAATTATGTAAAGAGTGTGATTATG	29042
Qy	871	CCCTACAGAAACCTTCAGAGTCTACCTCCCTATCCAGGATCCCC--GACTCCTTCCCC	928
Db	29043	CCCTACAGAAACCTTCAGAGTCTACCTCCCTACCCAGCATCCCCAGACTCCTTCCCC	29102
Qy	929	AMTAAATAGGACCCCTTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAA	988
Db	29103	AAATAATAGGACCCCTTCAACCCAAACGGTCCAAAAGGAGATAGACAAAAGGGTAAA	29162
Qy	989	CAGTGAAACAAAGAGTGCCAATATTCCCCCAATTATGACCCCTCCCAAGCAGTGGAGGAA	1048
Db	29163	CAACTAACCAAGAATGCCAATATTCCCGGATTATGCCCCCTCC---AAGCGGTGGAG	29218
Qy	1049	GAGAAATTCGGCCACGAGAGTGCAATGTCGCTTTTCTCCAGACTTAAAGCAATAA	1108
Db	29219	GAGAAATTCGGCCACGAGAGTGCAATGTCGCTTTTCTCTCAGACTT---TAAATTA	29274
Qy	1109	AAACAGACTTAGTAAATTTCTCAGATAATCCCTGATGGCTATATTGRTGTTTTCACAGGGT	1168
Db	29275	AAATAGACCTAGTAAATTTCTCAGATAACCTTAATGGCTATATTGATGTTTACAGGTT	29334
Qy	1169	TAGGACAAATCTTTGATCTGACATGGAGAGATATATATGTCATGCTCTAAATCAGACATA	1228
Db	29335	TAGGACAAATCTTTGATCTGATATGGAGAGATATA-ATGTTACTGCTAAATCAGACATA	29393
Qy	1229	ACCCCAATGAGAGAAGTCCACCAATACTGCAGCTGAGTGTTGGCGATCTCTGGTAT	1288
Db	29394	ACCCCAATGAGAGAAGTGCGCCGTAACTGCAGCTGAGTGTTGGCGATCTCTGGTAT	29453
Qy	1289	CTCAGTCAGGTCAAATGGATANGGATGACAAAGAGGAAAGANAATGATTTCCCCACAGGC	1348
Db	29454	CTCAGTCAGTCAATG---ATAGGTCGACACAGAGGAAAGAGACGATTTCCCAACGGC	29510
Qy	1349	CAGCARGCAGTTCACGCTCTASACCTCATTTGGGACACAGAAATCAGTAACATGGGAGA	1408
Db	29511	CAGCAGGCAGTTCACGCTGAGACCTCAGTGG--ACACAGATCAGAACAT-GGAGA	29566
Qy	1409	TTGCTGCTGCAGACATTTCTTAACTTGCTGCTASAGACTTAAGAAACCTASGAGAA	1468
Db	29567	TTGCTGCGCAGACATTTGCTTAACTTGGCTGCTAGAGGACTTAAGAAACCTAGAAAG-A	29625
Qy	1469	ARTCTAYGAAATTACTCAATGATGCCACATACACAGGGGAGGAGAAATCTCTACT	1528
Db	-29626	AGCTGTGAGTTTTCATGATGTCCTATTAACACA-GGGAAGAGAGAAATCTCTACC	29684
Qy	1529	GCCTTTCTGGAGACCTAAGGGAGGCAITGAGGAAGCGTGCCTCT--CTGTCACTGACT	1586
Db	29685	GCCTTTCTGGAGTGACTAACGGAGGCAITGAGGAAGCATACCTCTCTGTCACTGACT	29744
Qy	1587	CTTCTGAAGGCCAATACTTAAGCGTAACTTATCACTCAGTCAGCTGCAGACATTA	1646
Db	29745	CTACTGAAGGCCAATACTTAAGAGTAAGTTTATCACTCAGTCAGCTACAGACATTA	29804
Qy	1647	-GAAAAAATTTCAAAAGTCTCCGCTAGCGCCGGAGCAAACTTAGAAACCTTATTCGA	1705
Db	29805	GAAAAAATTTCAAAAGTCTCCCTTAGGCCCGGAACTTAGAAACCTTATTCGA	29864
Qy	1706	TGGCAACCTCGGTTTTTATTAATAGATCAGGAGGAGCAGCGGAAACAGGACAAACGGG	1765
Db	29865	TGGCAACCTCGGTTTTTATTAATAGATCAGGATGAGCAGGAGGAGTGGGACAAATGGG	29924
Qy	1766	ATTAATAAAGGCGACCGCTTTAGTCATGACCTCAGGCAAGTGGACTTTGGAGGCTCT	1825
Db	29925	A-TAAAAAAGGCGACCGCTTTAGTCATGCGCCTCAGGAGCGGACTTTGGAGGCACT	29983
Qy	1826	GAAAAAGGAAAGCTGGGCAATTTGAATGCTTAATAGGCTTTCAGTGGCGGCTCTA	1885
Db	29984	GAAAAAGGAAAGCTAGGCAATCAAAATGCTTAATAGGTTTGTCTTCAGTGGCGGCTCTA	30043
Qy	1886	CAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAAGCCGCCCTTCGTCATGCCCC	1945



US-09-949-016-5640

Query Match		29.4%;	Score 2204;	DB 3;	Length 2763;
Best Local Similarity		98.4%;	Pred. No. 0;		
Matches 2194;		Conservative 35;	Mismatches 0;	Indels 1;	Gaps 1;
QY	5352	GATACATCTCTGGGAGGACCTTACCCAGTCATTTTATYATACCCCAACTGGGGTTAAAGTG	5411		
DB	535	GATACATCTCTGGGAGGACCTTACCCAGTCATTTTATCTACCCCAACTGGGGTTAAAGTG	594		
QY	5412	GCTGGAGTGAGTCTTTGGATACATCACTTTGAGTCAAAATCTCTGGATCTGCGCAAGGAA	5471		
DB	595	GCTGGAGTGAGTCTTTGGATACATCACTTTGAGTCAAAATCTCTGGATCTGCGCAAGGAA	654		
QY	5472	CCTGAAATCCAGGACACCGTAGCTATCTCTGTGAACCTCTAGAGATTTGCGCCTG	5531		
DB	655	CCTGAAATCCAGGACACCGTAGCTATCTCTGTGAACCTCTAGAGATTTGCGCCTG	714		
QY	5532	CTCTTCAAAACAAACAGGAGGAAAGTAATAAATACTAAATCCCCATGSGCCTCC	5591		
DB	715	CTCTTCAAAACAAACAGGAGGAAAGTAATAAATACTAAATCCCCATGSGCCTCC	773		
QY	5592	TTATCATATTTTCTCTKTAAGTTSTTTTACCTSTTTTCACTCTCACTGCAACCCCTCC	5651		
DB	774	TTATCATATTTTCTCTKTAAGTTSTTTTACCTSTTTTCACTCTCACTGCAACCCCTCC	833		
QY	5652	ATGCGCTGTATGACCACTAGCTCCCTTACCCAGAGTTTCTATGGAGATGAGCGTCC	5711		
DB	834	ATGCGCTGTATGACCACTAGCTCCCTTACCCAGAGTTTCTATGGAGATGAGCGTCC	893		
QY	5712	CGGAAATATTTGATGCCCATCTGTATAGGAGTCTTTSTAAGGAAACCCCACTTCACTGC	5771		
DB	894	CGGAAATATTTGATGCCCATCTGTATAGGAGTCTTTSTAAGGAAACCCCACTTCACTGC	953		
QY	5772	CCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTTTTGTGATGCATGCAAAATAC	5831		
DB	954	CCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTTTTGTGATGCATGCAAAATAC	1013		
QY	5832	TCATTTTGGACAGGAAATGATTAATCTCTAGTTGTCTGGAGACTTGGAGTCACTGT	5891		
DB	1014	TCATTTTGGACAGGAAATGATTAATCTCTAGTTGTCTGGAGACTTGGAGTCACTGT	1073		
QY	5892	CTGTTGGACTTACTTACCACCAACTGATGTCTGATGGGGTGGAGTTCAAGATCAGGC	5951		
DB	1074	CTGTTGGACTTACTTACCACCAACTGATGTCTGATGGGGTGGAGTTCAAGATCAGGC	1133		
QY	5952	AAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCCGGGTACATGGCACTCTAG	6011		
DB	1134	AAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCCGGGTACATGGCACTCTAG	1193		
QY	6012	CCCCACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCCATCTCGCT	6071		
DB	1194	CCCCACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCCATCTCGCT	1253		
QY	6072	GCTAGCCCTTTTATACACCCCTCACTGGGTCCATGAGGTCTCGGCCCAAAACCCCTAC	6131		
DB	1254	GCTAGCCCTTTTATATACACCCCTCACTGGGTCCATGAGGTCTCGGCCCAAAACCCCTAC	1313		
QY	6132	TAACTGTTGGATATGCTCCCTCGAACTTTCARGGCATATGTTTCAATCCCTGTACCTGA	6191		
DB	1314	TAACTGTTGGATATGCTCCCTCGAACTTTCARGGCATATGTTTCAATCCCTGTACCTGA	1373		
QY	6192	ACAATGGAAACAATTCAGACAGAAATAAACCACTTCCGTTTTTAGTAGGACCTCTTGT	6251		
DB	1374	ACAATGGAAACAATTCAGACAGAAATAAACCACTTCCGTTTTTAGTAGGACCTCTTGT	1433		
QY	6252	TTCCAAATSTGGAATAAACCCATACCTCAAACTCTCACTGTGTAAATTTAGCAATAC	6311		
DB	1434	TTCCAAATCTGGAATAAACCCATACCTCAAACTCTCACTGTGTAAATTTAGCAATAC	1493		
QY	6312	ATACACAACCACTCCCAATGATCAGGTGGGTACTCTCCCAACCAATAGTCTGCT	6371		
DB	1494	ATACACAACCACTCCCAATGATCAGGTGGGTACTCTCTCCCAACCAATAGTCTGCT	1553		

QY	6372	ACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAAATGGCTCTTC	6431		
DB	1554	ACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAAATGGCTCTTC	1613		
QY	6432	AGAATCATGTGCTTCTCTCATTTAGTGGCCCAATGRCCTATCAGCTGAAACAAGA	6491		
DB	1614	AGAATCATGTGCTTCTCTCATTTAGTGGCCCAATGRCCTATCAGCTGAAACAAGA	1673		
QY	6492	TTTATACAGTTATGTATATCTTAAGCCCGCAACAAAGAGTACCCATTTCTTCTTTTGT	6551		
DB	1674	TTTATACAGTTATGTATATCTTAAGCCCGCAACAAAGAGTACCCATTTCTTCTTTTGT	1733		
QY	6552	TATAGGACGAGAGTGTAGTGCATAGTACTGCGATTGGCGTATCACAACCTCTAC	6611		
DB	1734	TATAGGACGAGAGTGTAGTGCATAGTACTGCGATTGGCGTATCACAACCTCTAC	1793		
QY	6612	TCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGGTGCGGACTC	6671		
DB	1794	TCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGGTGCGGACTC	1853		
QY	6672	CCTGCTCACCTTSCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTCRAAATCGAAG	6731		
DB	1854	CCTGCTCACCTTSCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTCRAAATCGAAG	1913		
QY	6732	AGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAAACCTGTTTATTTTAAAGGGAAGAAATG	6791		
DB	1914	AGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAAACCTGTTTATTTTAAAGGGAAGAAATG	1973		
QY	6792	CTGTTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAAATTCGAGATCG	6851		
DB	1974	CTGTTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAAATTCGAGATCG	2033		
QY	6852	AATACAAAGTAKAGCAGAGAGCTTCGAAACACTGACCCCTGGGGCTCCTCAGCCCATG	6911		
DB	2034	AATACAAAGTAKAGCAGAGAGCTTCGAAACACTGACCCCTGGGGCTCCTCAGCCCATG	2093		
QY	6912	GATGCCCTGGAATCTCCCTCTTAGGACCTCTAGCAGCTATAATATTTGCTACTCTCTT	6971		
DB	2094	GATGCCCTGGAATCTCCCTCTTAGGACCTCTAGCAGCTATAATATTTGCTACTCTCTT	2153		
QY	6972	TGGACCTGTATCTTTTAACTCTTTTAACTTTGTTTAACTTTGTTTAACTTTGTTTAA	7031		
DB	2154	TGGACCTGTATCTTTTAACTCTTTTAACTTTGTTTAACTTTGTTTAACTTTGTTTAA	2213		
QY	7032	ACTCAAAATGGAGCCCAAGATGACGCTCAAGACTAAGATCTACCGCAGACCCCTGACCG	7091		
DB	2214	ACTCAAAATGGAGCCCAAGATGACGCTCAAGACTAAGATCTACCGCAGACCCCTGACCG	2273		
QY	7092	GCTGTAGCCCAACGATCTGATTTAATGACATCAAAAGGACCCCTCTCTGAGGAATCTC	7151		
DB	2274	GCTGTAGCCCAACGATCTGATTTAATGACATCAAAAGGACCCCTCTCTGAGGAATCTC	2333		
QY	7152	AGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAGAGATTAGAGCGTGTGCGGC	7211		
DB	2334	AGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAGAGATTAGAGCGTGTGCGGC	2393		
QY	7212	AACCTCCCAACAGCAGCTTTAGTGTTCCTGTTGATGCGGACCTGAGAGCAGGACTAG	7271		
DB	2394	AACCTCCCAACAGCAGCTTTAGTGTTCCTGTTGATGCGGACCTGAGAGCAGGACTAG	2453		
QY	7272	CTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGTGTGGAAAGGTGACCAATCCAC	7331		
DB	2454	CTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGTGTGGAAAGGTGACCAATCCAC	2513		
QY	7332	CTTTAAACACGGGGCTTGCAACTTAGYTCAACCTGACCAATTCAGAGAGCTCACTAAAT	7391		
DB	2514	CTTTAAACACGGGGCTTGCAACTTAGYTCAACCTGACCAATTCAGAGAGCTCACTAAAT	2573		
QY	7392	GCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCAATCATTTATTCMTGAGGACACA	7451		
DB	2574	GCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCAATCATTTATTCMTGAGGACACA	2633		







Db 2641 CTACTATGCCGAGCAATCACTGGAAGGTGCACTGCCCCAGAGGACAAAGTTTCTTAGG 2700  
Qy 2003 TCAGAGCCACTAACAGAGATGATCCAGAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCA 2062  
Db 2701 TTGGAAGCCCCCAACTGATGATCCAAACAGGACTGAGGGTCCCGAGGCAAGCGCCA 2760  
Qy 2063 TCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGAACCATTTAGAGGCCAGGAA--- 2119  
Db 2761 GCTCATGTCACTCCCTACTGAGCCCAAGGTATGTTTAAACCATTTAGAGGCCAGGAAAT 2820  
Qy 2120 GGTGTCTCTGGAACACTGGTGGGTCTTCTTAGTCTTACTCTTCTGTCGCCGACAACTG 2179  
Db 2821 GACTTCTCTGGACACTGGCAGCTCTTCTCAGTGTCTTCTCTCTCTCTCTCTCTCTCTCT 2880  
Qy 2180 TCTTCCAGATCTGTCACTATTCTGAGGGGTGCTNWAAGAGCGGAGTCACTAGATATTT 2239  
Db 2881 TCTCTCAAGGTCTGTATACATCC---GAGGAATCCTGGGGCAGCTGTAAACCAAGT-GTTT 2936  
Qy 2240 YTCCAGCCACTAAGTTATGAACCTGGGGAGCTTTATTTCTTTTTCACATGCTTTCTAATTA 2299  
Db 2937 CTCCACCTCTCAGTTGT-AATTGGGAGACTTTGCTCTTTTCTGTATGCTTTTCTTGTGA 2995  
Qy 2300 TGCTTGAAGCCCACTACCTTGTGTAGGAGAGACATTTAGCAAAAGCAGGGGCCATTA 2359  
Db 2996 TGCTGAAAGTCCACACCTTTATAGGAGATATAT--TAGCCAAAGCTGGAGCTATTA 3053  
Qy 2360 TACACCTGAACATAGGAAAGGAACACCCGTTTGTGTGTCCTGCTTGTAGGAGGAAT 2419  
Db 3054 TCTACATGAATATGGGAAACAAGTTACCCATTTGTTGT-CCCTTACTTGGAGGGAAT 3112  
Qy 2420 AATCCTGAAGTCTGGGCAACAGAGGACATATGGAGGAGCCCAAGATGCCGTCCTGT 2479  
Db 3113 AACCTTGAAGTCTGGGCAATTTGGAGGACAATTTGGAGGGCAAAA-AATGCCACCCAGT 3171  
Qy 2480 TCAAGTTAACTAAAGGATTCACCTTCTTCTTACCAAGCAGTACCCCTCAGACC 2539  
Db 3172 CCAATCAGCTAAAGACTCCACCTTTCTTCTTATCAAGGCAATATCCCTTAAGGCC 3231  
Qy 2540 CAAGGCCCAACAGGATTCCAAAGATTTGAAGGACTTAAAGCCCAAGGCTTAGTAA 2599  
Db 3232 TGAAGCTCTTAAAGGATACAGGATATTTGAACATTTAAAGCTCAAGCTTAGTAAG 3291  
Qy 2600 ACCATGCATACCTTCCTGAGTAATTCCTGAGTGGATTTAGGAGGACAGAAACCCAGT 2659  
Db 3292 GAAATGCAGAGTCCCTGCAACACCC------AATTTAGGAGTACAAAAACCAATG 3344  
Qy 2660 GACAGTGGAGGTTAGTGAAGATCTCAGGATTTCAATGGAGGCGCTTGTCTTTTATA 2719  
Db 3345 GTCACCTGGA-GACTAGTGCAGATCTTAGACTCATCAAT-GAGGCAAGTAATTCCTCTATA 3402  
Qy 2720 CCAGCTGTACCTAGCCCTTATACCTGTGMYTTCCCAATATACAGAGGAAGCAGATGGTT 2779  
Db 3403 TCCAGTTGTACCCCAACCCCTATACCTGCTTTCTCAANTACAGAGAGGAGCAATGGTT 3462  
Qy 2780 TACASTCTGGACCTTMAAGATGCCCTTCTTGTGATCCCTGTATCATCTGACTCTCAAT 2839  
Db 3463 CATGGTTCTGGACCTCAAGGATGCCCTTCTGTTTCCC-CTGCACTCTGACTCCAGTT 3521  
Qy 2840 CTGTTTGTCTTTGAAGATCTTCAACCCACATCTCACTCACTCGACTTCTTTTACC 2899  
Db 3522 TCTGTTTGTCTTTGAGATCCCAAGACACACGTCCTCACTTACATGATGCTTGGC 3581  
Qy 2900 CCAAGGTTTCAGGGATAGYCCCATCTATTGTCAGGCAATAGCCCAAGACTTGAGYCA 2959  
Db 3582 ACAAGGTTTAGGGATAGCCCTCACCTGTTGTGAGGCACTGGCCCAAGATCAGGCCA 3641  
Qy 2960 RTWTCAATCTGGACACTTCTTGTCTTGTGTAGTGGATGATTTACTTTTGTGCTTAC 3019  
Db 3642 CTCTCAAGTCCAGGCACCTTGTGCTCTCAGTATGTGGATGATTTACTTTTGTGCTACCAG 3701  
Qy 3020 TTCAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTAAATTTCTGCVACCTGTGG 3079  
Db 3702 TTCAGAAAGCTCTATGCCAGCAGGCTACTCTAGATCTCTTGAATCTTCTACCTAATCAAG 3761

Qy 3080 CTACAWGGTTTCAAACCSARAGCTCARTCTCTCTCACAGCAGTTTAAATACTTAGGRC 3139  
Db 3762 GTACAAAGCATCTAGGTCAAAGGTGCAGCTTTCTTTACAGCAGGCTAAATATCTAGGCT 3821  
Qy 3140 ABATTTATCCAAAGGACCAAGGCTCAGTGAAGAAATATCCAGCCTTACTTGGCTTA 3199  
Db 3822 AATCTTAGCCAGAGGACCAAGGCTCAGCAAGGAAATGAATACAGCCTTACTTGGCTTA 3881  
Qy 3200 TCTCATCTCAAAACCTTAAAGCAACTAAGRRRTTCTTGGCTTAAYAGGTTTCTGCG 3259  
Db 3882 TCTTGGCTTAAGACATTAACAGTTGCAGGGTCTTGGAAATCACGGCTTTTGGCG 3941  
Qy 3260 AAWATGGAATCCCAAGGTTGGCAAATAGCCAGGYCATTAATAACASTAATTAAGGAAA 3319  
Db 3942 ACTATGGA-TCCCTGATACAGTGAATAGTCAAGGCCCTCCCATCTCTAATCAAGGAGA 4000  
Qy 3320 CTCAGAAAGCAATACCCATTTARTTAAGATGGAYANCTGAAGYMRBAAGTGGCTTCCAG 3379  
Db 4001 CCAGAGGGCAATACCTTATCTAGTAGAATGGTAAACAGGGGCAAGAACAGCCCTTCAAAA 4060  
Qy 3380 CCCTTAAAGAGCCCTTAAACCAAGYCCAGGTGTTAAAGYTTGCCCAACRGGGCAAGCTT 3439  
Db 4061 CCTTAAAGCAGGCCCTGTA---CAAGTTCAGCTTTAAGCTTCCAC-----AACTT 4111  
Qy 3440 TTSTTATATRTCACAGAAAAAACAAGAAAYAGCTCTRGAGTCTTACACAGTCCRAG 3499  
Db 4112 CTCTTATATGTACAG-AGAGAGCAGGATAGTCTTGGAGTCTTACTTACAGACTCTTG 4170  
Qy 3500 GGNAGAGCTTGCACCCYRTGGCBYACCTGASTAAGGAAATATGATGTAGTGGCAAGGTT 3559  
Db 4171 GGCAACCCCAACACAGTGGCATACCTAAGTAAGGAAATTTGATGTAGTACGAAAGGCT 4230  
Qy 3560 GRCTCATTTTAYGGGTAGTGGGAGTGGAGTGTATGATGTATCTGAAGCAGTTTAAAA 3619  
Db 4231 GGGCTCAGCTTTAAGGGTAGTTGAGCACTGGCCATCTTCATGTCAAGAGGCTTATCAAAA 4290  
Qy 3620 TAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCACTG 3679  
Db 4291 TAATACAGGAA-AGATCTCAGTCTGCTGACTACTCATGTTGTAAATGGCATACTAGTG 4349  
Qy 3680 CTAAGGAGACTTGGCTGTGAGCAACAGTGTCTTAAATTCAGGCTCTATTAATCTTG 3739  
Db 4350 CCAAGGAAATTTATGGCTATCAGACTACACCTACTTAGATACAGGCACTACTCTCTG 4409  
Qy 3740 AARGGCACTGTGCRACCTGCACTTGTGCACTTAAACCAGYCNCACTTCTTCTCCAG 3799  
Db 4410 AGGACCACTGTCTCAATATCTCATGTGTGGCCCTCAACCCCTGCACTTTTCTCCAG 4469  
Qy 3800 ACAATGAAGAAAGATARAAYATACTGTCAACAAATAATTTCTCAAACTATGCCACTC 3859  
Db 4470 AGAATGGGAAACCAATTTAGCATGACTGCCAACAAATTTACAGTCCAGACTTATGCCACCC 4529  
Qy 3860 GAGGGAGCTTGTAGAGTTCCTTGTGATGATCCGACCTTCAACTTGTATATCTGATGA 3919  
Db 4530 AAGATGATCTCTTAGAGTCCCTTAGCTTAATCTCTGACCTT-AACTATATACCGATGGA 4588  
Qy 3920 AGTTCCTTGTAGAAAAGGACTTCGAAAAGYGGGTATGCACTGTGCTGATGATATGA 3979  
Db 4589 AGTTCATTTGCAAGAAATGGGATATGAAGGGCAGGTTGTGACATAGTTAGTATGTAAC 4648  
Qy 3980 ATATYTTGAAATTAATCCCTCACTCCAGGAACTAGTGTCTGCTGCAAGCACTAATAGCC 4039  
Db 4649 GTACCTGAAAGTAAAGCTCTTCCCGCAGGACAGCACCCAGCTTAACAGACTAGTGGCA 4708  
Qy 4040 YTCATYKGGGCACTAGAATTAGGAGAGAAAAAGGGYAAATATATATATACAGACTCTRAR 4099  
Db 4709 CTTACTGAGCTTAGAACTGGGAAAGGAAAGAAATAAATGTTGTATACAGATACGAAG 4768  
Qy 4100 TATGCTYACCTAGTCTNTCCATGCCCATGMRGCAATATGSARAGAAAGGAAATCTTAACT 4159  
Db 4769 TATGCTTATCTAATCCCTCATGCCATGCTGCAATATGGAAGAAAGGAGGAGTTCTTAACC 4828

QY 4160 TCYAGRGAACACCTTATCAMACATCAGGAGCCATTGAGATTTATTATTCGCGTACAG 4219  
 Db 4829 TCCGGGGTACCCCATTAATATGACCAAGGAGATTTATGGAGTTATTCATGCTACGCA 4888  
 QY 4220 AAACCTATAGAGGTGCGAGTCTTACACTGCTGCGGTCTATCANAAAGGAAAGRAAGGAA 4279  
 Db 4889 AAGCCCAAGGAGTGGAGTCTTACACTGCGCAAGCCATCAGAAAGTGAAGGAGAAAG 4948  
 QY 4280 ATASAGRGAAYTGCAAGCAKATATTAAGAGMAAAGAGCTGCAAGGAGGACCCCTCA 4339  
 Db 4949 GCAGAAGGAAACCGTCAGGAGATCTGAGGCAAAATTTGCTGCGAGGTGGATCTCCCA 5008  
 QY 4340 TTAGAAATGCTTATTAACCTTCCCTTAGTATAGGTTAATCCCTCCGCGGAACCAAGCCC 4399  
 Db 5009 TTAGAAATACCTATGGAAGGACCTTGGAAATGGAACAAACCCCTCCAAGAGATTAAGCCC 5068  
 QY 4400 CAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGG---CAGTTTCTCCCTCG 4456  
 Db 5069 CAGTATTCGCCAAATGAAACAGAAATGGGAGTCTCATGGGGCATAGTTTCTCCCTCA 5128  
 QY 4457 GGACGGTTAGCCACTGAAGAGGAAATATCTTTTTCCTGCAACTATCAATGGAAATTA 4516  
 Db 5129 GGGTGGTTAAACGACAAAGGAAAGAGTACTTATACCGAAGCCAGCCAGTGGAAATA 5188  
 QY 4517 CTTAAACCCCTTCATCAACCTTTTCACTTAGGCATGATAGCAGCCATCAGATGGCCAA 4576  
 Db 5189 CTTAAACCCCTCCCAACCTTTTCAATATGGGTATGAAACACATCAATAGGCCACA 5248  
 QY 4577 TCATATTTACTGACACCGCTTTTCAAAACTATCAAGCARATATCAGGCGCTGTGAA 4636  
 Db 5249 TCTCTATTACAGGGCCAAATCTACTCTGGGCTTCCACAGGTAGTCAAACTGTGAG 5308  
 QY 4637 KTGTCGCAPARAAATAATCCCTGCTGTATGCCAAGCTCTTTCAGGAPAAACAAARACA 4696  
 Db 5309 GTGTGCAAGGAATAATCCCTGGTCCATTGTAAGGCCCATTTGGGGGAACAAAGAATA 5368  
 QY 4697 GGCCATTACCTGRRARARACTGGCAACTGATTTTACCCACAA--GCCCAACCTCAGG 4754  
 Db 5369 GGTCTACTATCTGGAGAGACTGGCAGTTAGACTTACCCATATGGCCCTAAATCAAGGG 5428  
 QY 4755 ATTTTCAGTATCTACTAGTCTGGGTARATATTTTACGGGTGGGCGARAGGCTTCCCTG 4814  
 Db 5429 ATTTCAATACTTGTGGTCTGTGTATACCTTTCAAAATGATAGAGCTTCCCTG 5488  
 QY 4815 TAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCATGAAATAATTTCCAGATT 4874  
 Db 5489 CAAGACAGAGAAGACTCAGGAAGTGATTGAAGTCTTAATTGATGAAATAATTTCTAGATT 5548  
 QY 4875 CGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTCCAGGCCACAGTAACCA 4934  
 Db 5549 TGGGCTTCCCTAAGGCTTACAGAGTGACAAATGGTCCAG--TTTAAAGGCCAATAACTCA 5607  
 QY 4935 GGGAGTATCCAGGCGTTAGGTATAGATATCACTTACACTGCGCTGAAAGGCCACAGTC 4994  
 Db 5608 GGGATGTCAGGCTGCTAGGATATCAGTATCACCTTCACTGCGCCGAGGCCAATC 5667  
 QY 4995 CTCAGGGAAGGTCGAGAAATGAATGAAYATCTAAAGGACATCTAAAGAACCAACCA 5054  
 Db 5668 CTCAGGGAAGGTCAGAGGCAAAATGAACACTCGAGGCATTAAGGAACTAAACACAAG 5727  
 QY 5055 GGAACCCACCTCATATGGCTGTCTGTGCTTATAGCTTTAAAGAGATCTGCACTT 5114  
 Db 5728 GAA----CATCTCCCATGGCTTCTTTTGGCCATGGCTTTGTGAGAATCCGAATTC 5783  
 QY 5115 TCCCAAAAGCAGGACTTAGGCCATACGAATGCTGTATGGAAGGCCCTTCAATACCAA 5174  
 Db 5784 TCTTCAAAATGGGCTCAGTCCATATGAATGCTGTATGACACATTTTCTCACAA 5843  
 QY 5175 TGACCTTGTGTGACCCCAAGACAG--CCAATTTAGTTGAGACATCCTCTCTTACCA 5233  
 Db 5844 TGACCTCTTACTTGTATGAGGAAAGACAAACTTTTCAAGATATAACTTCTTTCGCA 5903  
 QY 5234 ATATCAACAGTTCTTAAACATTACAGGAACCTTATCCCTGAGAGAGGGAAGAACT 5293

Db 5904 ATATCAACAAACCTTAARAACTACCTGAAGGATGTACAGAGAAAGGGAACAGATT 5963  
 QY 5294 ATTCACCCMWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTCCTCATCCCTAGA 5353  
 Db 5964 GTTTCAACAGGAGATCTAGTGGTGGTCAATCTCTCCCTCTACCTCCCATCTATGGA 6023  
 QY 5354 TACATCTGGGAAGGACCTTACCAGTCATTTTATYATACCCCAACTGCGGTTAAAGTGC 5413  
 Db 6024 CTCTCTGGGAAGGGTCATCTCAATAATCTCTTACCCACACTGCGATTAAGGTGT 6083  
 QY 5414 TGGAGTGAGTCTTGGATACATCAGCTTGAGTCAATCTCTGGATCTGCAAGGAAC 5473  
 Db 6084 AGGAGTGAATCTTGGATTACACACCCGAGTAAATTTTGGACATCCCTGAGGAAC 6143  
 QY 5474 TGAATAATCCAGGAG 5487  
 Db 6144 TGTGAGACCATCAG 6157

RESULT 15  
 US-09-854-867-26  
 ; Sequence 26, Application US/09854867  
 ; Patent No. 7014997  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOAN, KNOLL H  
 ; APPLICANT: ROGAN, PETER K  
 ; FILE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
 ; CURRENT APPLICATION NUMBER: US/09/854,867  
 ; CURRENT FILING DATE: 2003-05-08  
 ; NUMBER OF SEQ ID NOS: 613  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 26  
 ; LENGTH: 8399  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: repeat\_region  
 ; LOCATION: (1)..(8399)  
 ; OTHER INFORMATION: henv9  
 US-09-854-867-26

Query Match 28.9%; Score 2168; DB 5; Length 8399;  
 Best Local Similarity 67.6%; Fred. No. 0;  
 Matches 3334; Conservative 143; Mismatches 1354; Indels 103; Gaps 34;

QY 605 TGATGGGAAACGTTCCCGCAAGACAAACCGCCCTTAAGACGTATTTCTGGARAAATTGGG 664  
 Db 1276 TGATAGGAAACACTCAGGCATCAACGCTCACCTTGAATGTATCTTAGCCATTGGG 1335  
 QY 665 AMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTGAGTGGCGCT 724  
 Db 1336 ACCAATTTGACCAAAACCTGAAAGAGAGGAGCAGCTCATTTTTTCTGCACTACAGCT 1395  
 QY 725 GGCCT-----CCTGAGGAGTATAAATTATA 752  
 Db 1396 GACCCATAATTTCTCTCTGATGGGAAATAATGCGCCCTGAGGAAAGTATAAATTACA 1455  
 QY 753 ACACCATCTTACAGCTAGACTTCTTTCT---AGAAAAAGGCAAAATGGAGTGAAGTGCCAT 809  
 Db 1456 ATACTATCTGAGGTTGACCTTTTCTGTAGAGGAGAGGTAAATGGAGTGAATACCAT 1515  
 QY 810 AAG----TACAACTTTCTTTTCAATTAAGAGACAACTCAATATATGTAATAAAGTGTAT 865  
 Db 1516 AAGTATTACAAGCTTTCTTTTCACTGAAGGAGATACACACTATGCAAGCTTGCAT 1575  
 QY 866 TTATCCCTTACAGGAGCTTTCAGAGTCTACTCCCTATCCAGCATCCCGACTCTTC 925  
 Db 1576 TTACATCCCAAGGAGGAGCTTTCAGCTTACCCCATATCTCTAGCTCTCTTACAGCTCTC 1635  
 QY 926 C--CAATTAATAGGAGCCCTTTCACACCCCAAAATGGTCCAAAGGAGATAGACAAAGG 983

Db 1636 CTTCTATTAAATGCAAGCCTCTCCAAATCTCCCTGCCAGAGGAAATAAGCAAGAA 1695  
Qy 984 GTAAACAGTGAACCAAGAGTGCACATATATCCCAATATATGACCCCTCCCAAGAGTGGG 1043  
Db 1696 ATCTCAAGAGGACCAAAACCCCTGGGCTATCGGTATATGTCCTCCCTT-CAAGCTGTAGG 1754  
Qy 1044 AGGAAGAGAAATTCGGCCAGCAGAGTGCATGTCVTTTTYYTCTCCAGACTTAAAGCA 1103  
Db 1755 GGTGTGAGAAATTTGGCCCAACCCAGGTACATGTCCTCTCTCTGTGATTTAAAGCA 1814  
Qy 1104 AATAAAAAACAGACTTAGGTAAATTTCTCAGATAATYCTGTATGGCTATATATGTTTACA 1163  
Db 1815 GATCAAGGCAGACTGGGGGAAGTTTTTCAGATGATCTGTATAGGCATAGATGTCTCTCCA 1874  
Qy 1164 AGGGTTAGGCAATTTCTTGATCTGACATGGAGAGATATATATGTCACCTGCTAAATCAGA 1223  
Db 1875 GGGTCTAGGGCAACCTTCGATCTCATCTTGGAGAGATGT-CATGCTATTTGTAGATCAAA 1933  
Qy 1224 CACTAACCCCAATGAGAGAGTGCACCACTACTGCAGCTCGAGGTGTGGGATCTCT 1283  
Db 1934 CCTGGCCTTAAATGAAGAATGTGGCTGTAGCTGCAGCTGAGAGTTTGGAGATACCT 1993  
Qy 1284 GGTATCTCAGTCAAGTCAATGTGATANGATGACACAGAGGAGAAAGANAATGATTCCTCCA 1343  
Db 1994 GGTATCTTAGTCAAGTAAATG---ATAGAATGACAGCCGGAAGAGATAAATTCCTTA 2050  
Qy 1344 CAGGCCAGCARGCAGTTCAGCTTASACCTCATTTGGGGACACAGAAATCAGTAAATCAG 1403  
Db 2051 CTGGCCAGCAGCACTCCCAAGTATGGATGCCCTCCCACTGGGACCTC---AATCAGATCAT 2106  
Qy 1404 GGAGATTTGTGCTGCAGACATTTGCTAACTTGTGTCTASAGGACTAAGGAAACTASG 1463  
Db 2107 GGGGACTGAGTCTGAACATCTGTTGACTGTGTTCTGGAAGGACTAAGGAAATTAGG 2166  
Qy 1464 AAGAAATCTAVGAATTAATCAATGATGTCACCATAAACACAGGGGAAGGAAGAAATC 1523  
Db 2167 AA-AAAGCCCATGATTTATCAATGATGTCACCATTAATCA-GGGAAGGAAGAAATC 2224  
Qy 1524 CTATGCTCTTTCTGGAGAGACTAAGGGAGGCAATGAGGAAGCGTGTCTCTGTACCTG 1583  
Db 2225 CTTCTGCTCTTTCAAGTGGCTTCAGGAGACCTTAAAGAAATATATCTCCACTGTCTACCTG 2284  
Qy 1584 ACTCTTCTAAGGCCCAACTAATCTTAAAGGTAAGTTTATCACTAGTCACTGCTGCAGACA 1643  
Db 2285 AATCACTCGAGGGTCAATTTGATTTCTAANAAGATAGTTTATTAACCAATCAGCAGCAGATA 2344  
Qy 1644 TTAG-AAAAAATCTCAAAAGTCTGCGTAGGCCCGGAGCAAACTTTAGAAACCTTATGCA 1702  
Db 2345 TCAGTAGAAGCTCCAAAAGCAAGCCCTGGGCCCTGNACAAATATTTAGGCAATTTAA 2404  
Qy 1703 ACTTGGCAACYTGGTTTTTATATAGAGATCAGGAGGAGCAGGCGGACAGCAAAAC 1762  
Db 2405 ACTTGGCAACCGTGTGTCTATATATAGGACCAAGAGGACAGGCCCAA---AGGAA 2460  
Qy 1763 GGGATTAATAAAGAGCCACCGTTTAGTCATGACCCCTCAGGCAAGTGGACTTTGAGGC 1822  
Db 2461 AGAGATCAGAAAAGGCCGCGCTTAGTCATGGCCCTCAGACCAACAAACTTTGGTGT 2520  
Qy 1823 TCTGAAAAGGGAAGCTGGGCAATTTGATCCCTAATAGGCTTGTCTCCAGTCCGCT 1882  
Db 2521 TCAGAGGACAGAAATAGAGAGGCCCAATCAGCTGGTAGGGCTTTGTTATCAGTGTGT 2580  
Qy 1883 CTACAGGACACTTTAAAAAAGATTTCTCAAGTAGAAGTAAGCCGCCCTTGGTCCATGC 1942  
Db 2581 TTACAAGGACACTTTAAAAAAGATTTGTCCAGTCAGAAACAAGCTGCCCTTATCTGTGT 2640  
Qy 1943 CCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAGGCTTTTGTAG 2002  
Db 2641 CTACTATGCCGAGGCAATCACTGGAAGGTGCTGCCCCAGAGGACAAAGTTTCTTAGG 2700  
Qy 2003 TCAGAGGCCACTTAAACAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGGCCA 2062  
Db 2701 TTGGAAGCCCCAACCTGATGATCCAAACACAGGACTGAGGGTGGCCAGGGCAAGGCCA 2760

Qy 2063 TCCATGTCATCACTCCCTCAAGAGCCTCGGTATGCTTGAACCATGAGGCCAGGAA--- 2119  
Db 2761 GGTCTATGTATCACTACCCTCACTGAGCCCAAGGATATGTTTAAACCATGAGGCCAGGAAAT 2820  
Qy 2120 GGTGTCTCTCGGACACTGGTGGGTCTTCTTAGTCTTACTCTCTCTGTCCCGACACTG 2179  
Db 2821 GACTTCTCTGACACTGGCAGTCTTCTCAGTCTCTATCTCTCTTCTTCTGATGACTG 2880  
Qy 2180 TCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGCGGCGAGTCACTAGATACTTT 2239  
Db 2881 TCTCAAGTCTGTATTACCATCC---GAGGAATCTCTGGGCGAGCTGTAAACAGT-GTTT 2936  
Qy 2240 YTCCAGACCACTAAGTATTGAACCTGGGAGCTTTATCTTTTTCACATGCTTTTCTAATTA 2299  
Db 2937 CTCCCACTCTCTCAGTTGT-AAATGGGAGACTTTGCTCTTCTGTATGCTTTCTTGTTA 2995  
Qy 2300 TGTCTGAAGCCCTCACTACTCTGTTAGGGAGAGACATTTAGCAAAAGCAGGGCCATTA 2359  
Db 2996 TGTCTGAAGTCCCACTCTTTATTAGGGAGATATAT---TAGCCAAAGCTGGAGCTATT 3053  
Qy 2360 TACCTGAACATAGGAGAGGAAACACCCGTTTGTGTGTTMCCCTGCTTGAGGAAGAAAT 2419  
Db 3054 TCTACATGAATATGGGAACAAGTTTACCATTTGTGT-CCCTTACTTGAGGAGGAAAT 3112  
Qy 2420 AATCTGAAGTCTGGGCAACAGAAACAATATGGAAGGAGGCAAGAAATGCGCGTCTGT 2479  
Db 3113 AACCTTGAAGTCTGGGCAATTTGAAGGACAAATTTGGAAGGCAAAA-AAATGCCACCAGT 3171  
Qy 2480 TCAAGTTAACTAAGGATTCACCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2539  
Db 3172 CCAAAATCAGGCTTAAAGACTCCACACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3231  
Qy 2540 CAAGGCCCAACAGGATTCCTTAAAGACTTCTTAAAGACTTAAAGGCCCAAGGCTTCTTAA 2599  
Db 3232 TGAAGCTCTTAAAGATTAACAGATATTTGTTAAACATTTTAAAGCTCAAGGCTTAGTAAG 3291  
Qy 2600 ACCATGCAATCTCCCTGCGAGTAAATTCGTAGTGGATTGAGGAGGACAGAAACCCAGTG 2659  
Db 3292 GAAATGACAGCTCTCTGCAACACCC---AATCTAGGAGTACAAAAACCAATG 3344  
Qy 2660 GACAGTGGAGGTTAGTGAAGATCTCAGGATTTCAATGAGGAGCGCTGTCTTTTATA 2719  
Db 3345 GTCACTGGA-GACTAGTGAAGATCTTAGACTCATCAAT-GAGGCAAGTAAATCTCTTATA 3402  
Qy 2720 CCAGCTGTACCTAGCCCTTACTGTGMYTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2779  
Db 3403 TCCAGTTGTACCCAAACCCCTATACCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3462  
Qy 2780 TACATCTCTGGAACCTTTAGGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2839  
Db 3463 CATGGTCTTGAGACTCAAGGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3521  
Qy 2840 CTTGTTGCTTTTGAAGATATCTTAAACCCATCTCAACTCACTGAGACTTTTATAC 2899  
Db 3522 TCTGTTGCTTTTGAAGATCTTCAAGACCAAGACCAAGTCCCACTTATAGTATGCTTGGC 3581  
Qy 2900 CCAAGGTTTCAAGGATAGVCCCATCTATTTGGCCAGGCAATAGCCCAAGACTTGAGYCA 2959  
Db 3582 ACAAGGTTTAGGATAGCCCTCACTGTGTTGGTCAAGGCACTGGGCCAAGACTTAGGCCA 3641  
Qy 2960 RTYMTCACTACCTGGAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3019  
Db 3642 CTTCTCAAGTCAAGGCACTTTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3701  
Qy 3020 TTCAGAAACCTTGTGCCATCAAGCAACCAAGCTCTTMAATTTCTCGYACCTGTGG 3079  
Db 3702 TTCAGAGGCTCATGCCAGGCTACTCTAGATCTCTTGAACCTTCTTCTTCTTCTTCTTCT 3761  
Qy 3080 CTACAGGTTTCCAAACASABAGCTCARTCTGTCTCAGCAGGCTTAAATCTTAGGCT 3139  
Db 3762 GTACAAGGCACTAGGTCAAGGTCAGCTTTGCTTTACAGCAGGCTTAAATCTTAGGCT 3821



```

Db      5964 GTTTCACCAAGGAGATCTAGTGTGGTCAAAATCTCTCCCTCTACCTCCCATCTATGGA 6023
QY      5354 TACATCCTGGGAAGGACCTTACCCAGTCAATTTTATYTAACCCAACTGCGGTTAAAGTGC 5413
Db      6024 CTCTCTGTGGGAAGGGTCATACTCAATAATCTCTTACCCACACTGCAGTTAAGGTGGT 6083
QY      5414 TGGAGTGGAGTCTTTGGATACATCACACTTTGAGTCAAAATCCTGGATCTGCCAAGGAACC 5473
Db      6084 AGGAGTGAATCTTGGATTCCACACCCCGAGTTAAATTTTGGACATCCCTCGAGGAACC 6143
QY      5474 TGAATATCCAGGAG 5487
Db      6144 TGTGAGACCATCAG 6157

```

Search completed: August 2, 2006, 02:04:52  
Job time : 1355 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:21:54 ; Search time 6527 Seconds  
(without alignments)  
8099.217 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7489.6  
Sequence: 1 caacaatgggataaacc.....tattaatcttgcaactgcr 7582

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 8:\*

1:	Geneseqn1980s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004as:*
13:	Geneseqn2004bs:*
14:	Geneseqn2005s:*
15:	Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7489.6	100.0	7582	3	AAA59215 Human end
2	7487.4	99.9	7582	2	AXX25665 Complete
3	6448	86.1	56093	6	ABL61744 Colon ade
4	6436	85.9	10499	3	ABN97929 Human ret
5	6124.6	81.8	8523	7	ADS30988 Human gen
6	6124.6	81.8	8523	7	ADP36376 HIRA geno
7	5108	68.2	9502	10	ADF59718 Human con
8	3738.4	49.9	6394	5	AAA59210 DNA encod
9	3431	45.8	4349	5	AAA59215 DNA encod
10	3309.6	44.2	7466	5	AAA59215 DNA encod
11	3049.8	40.7	22436	9	ADA02882 Human PAP
12	3049.8	40.7	22436	10	ADB72620 Human PAP
13	3049.8	40.7	22436	10	ADC85361 Mouse Pap
14	3049.8	40.7	22436	12	ADM74477 Human car
15	3049.8	40.7	22436	14	ADZ12709 Human can
16	2948	39.4	139573	10	ADH58564 Human Na+
17	2948	39.4	156416	13	ADB32817 Human can
18	2932	39.1	5154	5	AAA59215 DNA encod

19	2908.2	38.8	3464	10	ADE09587	Adc09587 Novel DNA
20	2889.6	38.6	2930	6	AAZ24195	AAZ24195 Human syn
21	2888.6	38.6	2946	2	AAZ77526	AAZ77526 Human sec
22	2888.6	38.6	2946	3	AAZ59468	AAZ59468 Human sec
23	2888.6	38.6	2946	10	ADC38776	ADC38776 Human con
24	2734.4	36.5	3831	5	AAZ71727	AAZ71727 DNA encod
25	2733	36.5	102200	15	ABF74640	ABF74640 Human pol
26	2733	36.5	102200	15	ABF74983	ABF74983 Human pol
27	2688.6	35.9	2938	2	AAZ25656	AAZ25656 Human end
28	2688.6	35.9	2938	3	AAA59206	AAA59206 Gag and p
29	2670.8	35.7	173564	13	ABD32953	ABD32953 Human can
30	2613.2	34.9	3372	2	AAZ25663	AAZ25663 Human end
31	2613.2	34.9	3372	3	AAA59213	AAA59213 Partial p
32	2550.4	34.1	2599	3	ABN97927	ABN97927 Human ret
33	2541.6	33.9	4535	5	AAZ76205	AAZ76205 DNA encod
34	2541.6	33.9	4535	10	ADC32196	ADC32196 Human nov
35	2368.4	31.6	8279	5	AAZ76474	AAZ76474 DNA encod
36	2282.4	30.5	2784	3	ABN97930	ABN97930 Human ret
37	2251.8	30.1	2372	2	AAZ25664	AAZ25664 Human end
38	2251.8	30.1	2372	3	AAA59214	AAA59214 R-U5 regi
39	2241	29.9	8294	5	AAZ84209	AAZ84209 DNA encod
40	2241	29.9	8294	10	ADF59732	ADF59732 Human con
41	2220.4	29.6	161334	11	ACN44334	ACN44334 Human gen
42	2216	29.6	46340	3	ABN97978	ABN97978 Human ret
43	2204.6	29.4	2782	2	AAZ25661	AAZ25661 Human end
44	2204.6	29.4	2782	3	AAA59211	AAA59211 5' non co
45	2204.6	29.4	2782	5	AAH20069	AAH20069 HERV-W en

## ALIGNMENTS

RESULT 1  
AAA59215  
ID AAA59215 standard; DNA; 7582 BP.  
XX  
AC AAA59215;

DT 07-NOV-2000 (first entry)

XX Human endogenous retrovirus W (HERV-W) sequence.

XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W; gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX Human endogenous retrovirus.

XX Location/Qualifiers

FT LTR 1..120

FT /tag= a

FT /note= "R of 5' LTR"

FT 121..575

FT /tag= b

FT /note= "U5 of 5' LTR"

FT 579..596

FT /tag= c

FT 5581..7194

FT /tag= d

FT /note= "ORF1 env538"

FT 7039..7194

FT /tag= e

FT /note= "ORF2 52 AA"

FT 7112..7255

FT /tag= f

FT /note= "ORF3 48 AA"

FT 7244..7254

FT /tag= g

FT /note= "polypurine tract"

FT 7256..7582

FT /tag= h

FT /note= "U3-R of 3' LTR"

FT 7563..7569

FT /tag= i

FT polyA\_signal



XX WO200043521-A2.  
 XX PD 27-JUL-2000.  
 XX 21-JAN-2000; 2000WO-FR000144.  
 XX PF 21-JAN-1999; 99FR-00000888.  
 XX PR 21-JAN-1999; 99FR-00000888.  
 XX PA (INMR ) BIO MERIEUX.  
 XX PI Paranhos-Baccala G, Mallet F, Voisset C;  
 XX DR WPI; 2000-499229/44.  
 XX PT New nucleic acid from human endogenous retrovirus, useful e.g. for  
 XX PT diagnosis of autoimmune disease and complications of pregnancy, contains  
 XX PT at least part of the gag gene.  
 XX PS Disclosure; Page 49-52; 53pp; French.  
 XX CC The present sequence represents an endogenous retrovirus, which is  
 CC associated with an autoimmune disease, and is integrated into the human  
 CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The  
 CC HERV-W retrovirus is associated with autoimmune disease, failure of  
 CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or  
 CC proteins derived from it, are useful for diagnosis of autoimmune disease  
 CC (specifically multiple sclerosis) and for monitoring pregnancy. The  
 CC nucleic acid fragments may also be used for in situ labelling of isolated  
 CC chromosomes, while the transcription product can be used to study or  
 CC monitor T cell proliferation in vitro  
 XX SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;  
 Query Match 100.0%; Score 7489.6; DB 3; Length 7582;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAACAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC 60  
 Db 1 CAACAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC 60  
 QY 61 CTTCCCTTTGTATGGAGCTGTTTCATGCTATTTTCACTCTATTAATCTTGCACACTGCA 120  
 Db 61 CTTCCCTTTGTATGGAGCTGTTTCATGCTATTTTCACTCTATTAATCTTGCACACTGCA 120  
 QY 121 CTTCTTCTGTCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180  
 Db 121 CTTCTTCTGTCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180  
 QY 181 TGTTTGCCACCAACCGCANACTGCGCTGACTCCCATCCCTCTGGATCCTGCAAGGTGTC 240  
 Db 181 TGTTTGCCACCAACCGCANACTGCGCTGACTCCCATCCCTCTGGATCCTGCAAGGTGTC 240  
 QY 241 CGCTGTGCTTGTATCCAGCGAGCGCCCAATTCGCGCTCCCAATTTGGGCTAAAGCTTGC 300  
 Db 241 CGCTGTGCTTGTATCCAGCGAGCGCCCAATTCGCGCTCCCAATTTGGGCTAAAGCTTGC 300  
 QY 301 CATTGTNCTGCAACGGCTAAAGTCCCTGGGTTTGTCTAATTCAGCTGAACACTANTCACT 360  
 Db 301 CATTGTNCTGCAACGGCTAAAGTCCCTGGGTTTGTCTAATTCAGCTGAACACTANTCACT 360  
 QY 361 GGGTTCATGTTCTTCTTGTGACCCAGCGGCTTCTAATAKAACTATACACTTACCA 420  
 Db 361 GGGTTCATGTTCTTCTTGTGACCCAGCGGCTTCTAATAKAACTATACACTTACCA 420  
 QY 421 TGGCCCCAGATTCATTCCTTGGATCCGTGAGGSCAAGCACTCCAGGTGAGCAATAC 480  
 Db 421 TGGCCCCAGATTCATTCCTTGGATCCGTGAGGSCAAGCACTCCAGGTGAGCAATAC 480  
 QY 481 GARGCTTGGCACCATCTTGGAGCGGCTGCTACCTCTTGGAGTGGTTTCAACCACTC 540  
 Db 481 GARGCTTGGCACCATCTTGGAGCGGCTGCTACCTCTTGGAGTGGTTTCAACCACTC 540

QY 541 TTGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGCRACCAACGAGCATCC 600  
 Db 541 TTGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGCRACCAACGAGCATCC 600  
 QY 601 MAAGTGATGGGAAACGTTCCCGCAAGACAAAAACGCCCTTAAGACGCTATTCTGGAAT 660  
 Db 601 MAAGTGATGGGAAACGTTCCCGCAAGACAAAAACGCCCTTAAGACGCTATTCTGGAAT 660  
 QY 661 TGGGACAAATTTGACCTCTGAGACACTAAGAAAGAAACGACTTATATTCTTCTGAGTCC 720  
 Db 661 TGGGACAAATTTGACCTCTGAGACACTAAGAAAGAAACGACTTATATTCTTCTGAGTCC 720  
 QY 721 GCCTGCACTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACACTTTTGG 780  
 Db 721 GCCTGCACTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACACTTTTGG 780  
 QY 781 TAGAAAAAGGCAAAATGGAGTGAAGTCCCAATAGTACAAACTTTCTTTTCAATTAAGACAA 840  
 Db 781 TAGAAAAAGGCAAAATGGAGTGAAGTCCCAATAGTACAAACTTTCTTTTCAATTAAGACAA 840  
 QY 841 CTCACAAATTTGTAATAAAGTGTGATTTATGCCCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900  
 Db 841 CTCACAAATTTGTAATAAAGTGTGATTTATGCCCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900  
 QY 901 CTATCCAGCATCCCGACTCTTCCCGAMYTATAAGGACCCCGCTTCAACCCCAATGG 960  
 Db 901 CTATCCAGCATCCCGACTCTTCCCGAMYTATAAGGACCCCGCTTCAACCCCAATGG 960  
 QY 961 TCCAAAAGGAGATAGACAAAGAGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCAAT 1020  
 Db 961 TCCAAAAGGAGATAGACAAAGAGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCAAT 1020  
 QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGSCCCAGCAGAGTGCATGTGCT 1080  
 Db 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGSCCCAGCAGAGTGCATGTGCT 1080  
 QY 1081 TTTTCTTCCAGACTTAAAGCAATTAACACAGACTTAGTAAATTTCTGATTAATTCCT 1140  
 Db 1081 TTTTCTTCCAGACTTAAAGCAATTAACACAGACTTAGTAAATTTCTGATTAATTCCT 1140  
 QY 1141 GATGGCTATATTGTTGTTTACAGGGTTAGGACAAATTTCTTGTATCTGACATGGAGAT 1200  
 Db 1141 GATGGCTATATTGTTGTTTACAGGGTTAGGACAAATTTCTTGTATCTGACATGGAGAT 1200  
 QY 1201 ATATATGCTACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACCACTAACTGC 1260  
 Db 1201 ATATATGCTACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACCACTAACTGC 1260  
 QY 1261 AGCTGAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATGATGACACACA 1320  
 Db 1261 AGCTGAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATGATGACACACA 1320  
 QY 1321 GAAGAAAGANAATGATTTCCCAAGCAGCAGCAGTCCAGTCTASACCCCTCATTTG 1380  
 Db 1321 GAAGAAAGANAATGATTTCCCAAGCAGCAGCAGTCCAGTCTASACCCCTCATTTG 1380  
 QY 1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTCTGCAGACATTTGCTAACTTGTGTGC 1440  
 Db 1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTCTGCAGACATTTGCTAACTTGTGTGC 1440  
 QY 1441 TASAAGGACTAAGGAAACCTASAGAAAGAAATCTATGAATTAATCTAATGATGTCACCAATA 1500  
 Db 1441 TASAAGGACTAAGGAAACCTASAGAAAGAAATCTATGAATTAATCTAATGATGTCACCAATA 1500  
 QY 1501 ACACAGGGAAGGAAAGAAATCTTACTGCTTTCTGGAGAGACTAAGCGGAGGCTTGCAG 1560  
 Db 1501 ACACAGGGAAGGAAAGAAATCTTACTGCTTTCTGGAGAGACTAAGCGGAGGCTTGCAG 1560  
 QY 1561 GAAGCGTCTCTCTGTGCTCACCTGACTCTTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620  
 Db 1561 GAAGCGTCTCTCTGTGCTCACCTGACTCTTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620

Qy 1621 TATCACTCAGTCAGTCAGACATTAGAAAAAAGCTTCAAAAGCTCGCGTAGGCCGGAG 1680  
Db 1621 TATCACTCAGTCAGTCAGACATTAGAAAAAAGCTTCAAAAGCTCGCGTAGGCCGGAG 1680  
Qy 1681 CAAAACCTTAGAAACCTTATGAACCTTGGCAACCTCGGTATTTTATATAGAGATCAGGAG 1740  
Db 1681 CAAAACCTTAGAAACCTTATGAACCTTGGCAACCTCGGTATTTTATATAGAGATCAGGAG 1740  
Qy 1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGACCAACCGCTTTAGTCATGACCTT 1800  
Db 1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGACCAACCGCTTTAGTCATGACCTT 1800  
Qy 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTTGGCAAAATTTGAATGCTCTAA 1860  
Db 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTTGGCAAAATTTGAATGCTCTAA 1860  
Qy 1861 TAGGGCTTGCTTCAGTGCGGTCTACAGGACACTTTTAAAGGATTTGTCCAAAGTAGAAG 1920  
Db 1861 TAGGGCTTGCTTCAGTGCGGTCTACAGGACACTTTTAAAGGATTTGTCCAAAGTAGAAG 1920  
Qy 1921 TAAGCGCGCCCTTCGTCCATGCGCCCTTATTTCAAGGGAATCACTGGAAAGCCCACTGCC 1980  
Db 1921 TAAGCGCGCCCTTCGTCCATGCGCCCTTATTTCAAGGGAATCACTGGAAAGCCCACTGCC 1980  
Qy 1981 CAGGGACAAAGGTCTTTTGAAGTCAGAGCCACTTAACCAAGATGATCCAGCAGGACTG 2040  
Db 1981 CAGGGACAAAGGTCTTTTGAAGTCAGAGCCACTTAACCAAGATGATCCAGCAGGACTG 2040  
Qy 2041 AGGGTGCTGGGCAAGCGCATCCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTT 2100  
Db 2041 AGGGTGCTGGGCAAGCGCATCCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTT 2100  
Qy 2101 GACCATTAGGGCCAGGAAGGTGTCTCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
Db 2101 GACCATTAGGGCCAGGAAGGTGTCTCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
Qy 2161 CTTCGTCCCGGACACTGTCTCCAGATCTGTCACTATCTCAGGGGTCCWTAAGAG 2220  
Db 2161 CTTCGTCCCGGACACTGTCTCCAGATCTGTCACTATCTCAGGGGTCCWTAAGAG 2220  
Qy 2221 GGCAGTCACTAGATCTTTTCCAGCCACTAAGTTATGAACCTGGGGAGCTTTATCTTT 2280  
Db 2221 GGCAGTCACTAGATCTTTTCCAGCCACTAAGTTATGAACCTGGGGAGCTTTATCTTT 2280  
Qy 2281 TCACATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTA 2340  
Qy 2341 GCAAAAGCAGGGCCATTATACCTGAAACATAGGAGGAACACCCGTTTGTGTGCC 2400  
Db 2341 GCAAAAGCAGGGCCATTATACCTGAAACATAGGAGGAACACCCGTTTGTGTGCC 2400  
Qy 2401 CTGCTTGAGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAAATATGGAGC 2460  
Db 2401 CTGCTTGAGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAAATATGGAGC 2460  
Qy 2461 CAAAGAAATCCCGTCTGTTCAAGTTAACTAAAGGATTCACCTCTCTTTCCCTACCAA 2520  
Db 2461 CAAAGAAATCCCGTCTGTTCAAGTTAACTAAAGGATTCACCTCTCTTTCCCTACCAA 2520  
Qy 2521 GGCAGTACCCCTCAGACCAAGGCCCAACAGAGATTCCAAAGATTTGAAGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCAAGGCCCAACAGAGATTCCAAAGATTTGAAGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTAGTAAACCATATCACTCCCTGAGTAATTCCTGAGTGGATTGAG 2640  
Db 2581 AAGCCCAAGGCTTAGTAAACCATATCACTCCCTGAGTAATTCCTGAGTGGATTGAG 2640  
Qy 2641 GAGGCACAGAAACCCAGTGGAGGGTGTAGTGAAGATCTCAGGATTTCAATGG 2700  
Db 2641 GAGGCACAGAAACCCAGTGGAGGGTGTAGTGAAGATCTCAGGATTTCAATGG 2700  
Qy 2701 AGGCGGTGTCCTTTTATACCGAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAATAC 2760

Db 2701 AGGCGGTGTCCTTTTATACCGAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAATAC 2760  
Qy 2761 CAGAGAAGCAGAGTGTGTTTACASTCCTCGACCTTMAAGATGCTTCTTCTGSCATCCCTG 2820  
Db 2761 CAGAGAAGCAGAGTGTGTTTACASTCCTCGACCTTMAAGATGCTTCTTCTGSCATCCCTG 2820  
Qy 2821 TACATCCTGACTCTCAATCTTGTGTTTGGCTTTTGAAGATCTTCAAAACCCARCATCTCAAC 2880  
Db 2821 TACATCCTGACTCTCAATCTTGTGTTTGGCTTTTGAAGATCTTCAAAACCCARCATCTCAAC 2880  
Qy 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
Db 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
Qy 2941 TAGCCCAAGACTGTGAGYCAATYMTCATCTGAGCACTCTTGTCTTCTCRGTAGTGTGATG 3000  
Db 2941 TAGCCCAAGACTGTGAGYCAATYMTCATCTGAGCACTCTTGTCTTCTCRGTAGTGTGATG 3000  
Qy 3001 ATTTACTTTTTRGCGCYRTTTCAGAAACCTTGTGCCATCAAGCCACCCAGCRCTCTTMA 3060  
Db 3001 ATTTACTTTTTRGCGCYRTTTCAGAAACCTTGTGCCATCAAGCCACCCAGCRCTCTTMA 3060  
Qy 3061 ATTTCTCTCGYACCTGTGCTACAGGTTTCCAAACSAARAGTCTARCTCTGTCTCACAGC 3120  
Db 3061 ATTTCTCTCGYACCTGTGCTACAGGTTTCCAAACSAARAGTCTARCTCTGTCTCACAGC 3120  
Qy 3121 AGGTTAAATACCTTAGGCTTAARATATCCAAAGGACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Db 3121 AGGTTAAATACCTTAGGCTTAARATATCCAAAGGACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Qy 3181 TCCAGCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGGRRTTCTTCTG 3240  
Db 3181 TCCAGCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGGRRTTCTTCTG 3240  
Qy 3241 GRTTAAAGGTTTCTGCCGAATGGAATTTCCCAGGTTGGCRAAAATAGCCAGGYCATTA 3300  
Db 3241 GRTTAAAGGTTTCTGCCGAATGGAATTTCCCAGGTTGGCRAAAATAGCCAGGYCATTA 3300  
Qy 3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTARTAGATGGAAMCTGAA 3360  
Db 3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTARTAGATGGAAMCTGAA 3360  
Qy 3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCAGYCCCAGTGTTAAGTT 3420  
Db 3361 GYMAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCAGYCCCAGTGTTAAGTT 3420  
Qy 3421 TGCCAAACRGGGCAAGACTTTTSTTVATAYRTCAAGAAAAAACAGRAAYAGCTCTRGA 3480  
Db 3421 TGCCAAACRGGGCAAGACTTTTSTTVATAYRTCAAGAAAAAACAGRAAYAGCTCTRGA 3480  
Qy 3481 GTCCCTTACACAGTCCRAGGAYGAGCTTGAACCCYRTGGCRYACCTGASTAGGAAYT 3540  
Db 3481 GTCCCTTACACAGTCCRAGGAYGAGCTTGAACCCYRTGGCRYACCTGASTAGGAAYT 3540  
Qy 3541 GATGTAGTGGCAAGGTTGRCYTCAATTTTAYGGTAGTGTGGCTGSCAGTACAGTYKTA 3600  
Db 3541 GATGTAGTGGCAAGGTTGRCYTCAATTTTAYGGTAGTGTGGCTGSCAGTACAGTYKTA 3600  
Qy 3601 GTATCTGAAGCAGTTTAAATAATACAGGGAGAGATCTTACTGTGTGGAATCTCATGAK 3660  
Db 3601 GTATCTGAAGCAGTTTAAATAATACAGGGAGAGATCTTACTGTGTGGAATCTCATGAK 3660  
Qy 3661 GTGAAYRCATACCTCCTCTAAAGGAGACTTGTGGCTGTGAGCAACCYGTTTACTTAA 3720  
Db 3661 GTGAAYRCATACCTCCTCTAAAGGAGACTTGTGGCTGTGAGCAACCYGTTTACTTAA 3720  
Qy 3721 TRTCAAGGCTCTATTACTTGAAGGCCAGTGTGCRACCTGTGACCTTGTGCAACTCTTAAC 3780  
Db 3721 TRTCAAGGCTCTATTACTTGAAGGCCAGTGTGCRACCTGTGACCTTGTGCAACTCTTAAC 3780  
Qy 3781 CCAGYCNCATTTCTTCCAGACAAATGAAGAAAGATARAAYATACTGTCAACAAATAAT 3840

Db 3781 CCAGYCNCAATTTCTTCCAGACAAATGAAGAAAGATARAAYATAAATCTGTCAACAARTAAAT 3840  
Qy 3841 TCTCAAACTATGCCACTCGAGGAGACCTTTTAGAGTTTCCYTTGACATGATCCYGAACCTT 3900  
Db 3841 TCTCAAACTATGCCACTCGAGGAGACCTTTTAGAGTTTCCYTTGACATGATCCYGAACCTT 3900  
Qy 3901 CAACCTTGATATCTAGTGAAGTTCCCTTTGTAGAAAAGAGACTTCCAAAAGYGGGTATGC 3960  
Db 3901 CAACCTTGATATCTAGTGAAGTTCCCTTTGTAGAAAAGAGACTTCCAAAAGYGGGTATGC 3960  
Qy 3961 AGTGTCTAGTGAATATGGAATATTTGAAAGTAATCCCTCTCACTCCAGGAACCTAGTGTCTYA 4020  
Db 3961 AGTGTCTAGTGAATATGGAATATTTGAAAGTAATCCCTCTCACTCCAGGAACCTAGTGTCTYA 4020  
Qy 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTAGGAGAAAGRAAAAAGGGYAAA 4080  
Db 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTAGGAGAAAGRAAAAAGGGYAAA 4080  
Qy 4081 TATATATACAGACTCTRTARTAGCTYACCTAGTONTCCATGCCATGMRGCAATATGSAR 4140  
Db 4081 TATATATACAGACTCTRTARTAGCTYACCTAGTONTCCATGCCATGMRGCAATATGSAR 4140  
Qy 4141 AGAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db 4141 AGAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGTGMAGTCTTTACCTGCGYGGGTATCA 4260  
Db 4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGTGMAGTCTTTACCTGCGYGGGTATCA 4260  
Qy 4261 NAAAGGAAGRAAGGGAAATASAGRAAYTGCCAGCAKATATTTGAAGMAAAGAGC 4320  
Db 4261 NAAAGGAAGRAAGGGAAATASAGRAAYTGCCAGCAKATATTTGAAGMAAAGAGC 4320  
Qy 4321 TCGAAGGAGGACCTTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Db 4321 TCGAAGGAGGACCTTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Qy 4381 CTTCCGGGAAACCAAGCCCACTACTAGCAGGAGAAACAGAAATGGGAAACCTCAAGAG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCACTACTAGCAGGAGAAACAGAAATGGGAAACCTCAAGAG 4440  
Qy 4441 CAGTTTTCTCCCTCGGAGGCTTAGCCACTCAAGAGGAAATACCTTTTGCCTGCAAC 4500  
Db 4441 CAGTTTTCTCCCTCGGAGGCTTAGCCACTCAAGAGGAAATACCTTTTGCCTGCAAC 4500  
Qy 4501 TATCCAAATGGAATTAATTAATTTACTTCAAAACCTTCACTAGGATCGATAGCAC 4560  
Db 4501 TATCCAAATGGAATTAATTAATTTACTTCAAAACCTTCACTAGGATCGATAGCAC 4560  
Qy 4561 CCATCARATGGCCAAATCAATTAATTTACTGGAACAGGCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGGCCAAATCAATTAATTTACTGGAACAGGCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAGGCTGTGAATGTGCCARAAATATCCCTGCTCTYATGCCAGCTCCTTC 4680  
Db 4621 AKTCAGGCTGTGAATGTGCCARAAATATCCCTGCTCTYATGCCAGCTCCTTC 4680  
Qy 4681 AGGARAACAAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db 4681 AGGARAACAAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Qy 4741 CCCAAACCTCAGGATTTTCAATTAATTTACTAGTCTGGGTARATACCTTTTCAAGGGTGGCA 4800  
Db 4741 CCCAAACCTCAGGATTTTCAATTAATTTACTAGTCTGGGTARATACCTTTTCAAGGGTGGCA 4800  
Qy 4801 RAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTATGAA 4860  
Db 4801 RAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTATGAA 4860  
Qy 4861 ATAAATTTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGACAATAGCCCTGTTTCCAG 4920  
Db 4861 ATAAATTTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGACAATAGCCCTGTTTCCAG 4920

Qy 4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
Db 4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAATGAATGAAYACTCAAAAGGACATCTA 5040  
Db 4981 TGAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAATGAATGAAYACTCAAAAGGACATCTA 5040  
Qy 5041 AAAAAGCAACCCAGGAAAACCCACCTCATGCGCTGTTCTGTGCTTATAGCCTTAAAA 5100  
Db 5041 AAAAAGCAACCCAGGAAAACCCACCTCATGCGCTGTTCTGTGCTTATAGCCTTAAAA 5100  
Qy 5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAG 5160  
Db 5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAG 5160  
Qy 5161 CCTTCTATAACCAATGACTTTGTGCTTTGACCAAGACAGCCAACTTAGTTCGAGACATCA 5220  
Db 5161 CCTTCTATAACCAATGACTTTGTGCTTTGACCAAGACAGCCAACTTAGTTCGAGACATCA 5220  
Qy 5221 CCTCTTAGCCAAATATCAAGAGTTCTTAAAAATTTAACAAGAAACCTTATCCCTGAGAAG 5280  
Db 5221 CCTCTTAGCCAAATATCAAGAGTTCTTAAAAATTTAACAAGAAACCTTATCCCTGAGAAG 5280  
Qy 5281 AGGGAAGAACTATTTCCACCMMGTGACATGTTAGTCAAGTCCCTTCVCTCTAAT 5340  
Db 5281 AGGGAAGAACTATTTCCACCMMGTGACATGTTAGTCAAGTCCCTTCVCTCTAAT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCTTACCAGTCAATTTTATYTAACCCCACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCTTACCAGTCAATTTTATYTAACCCCACTG 5400  
Qy 5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGGAATACATCACCTTGAGTCAAAATCTGGATAC 5460  
Db 5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGGAATACATCACCTTGAGTCAAAATCTGGATAC 5460  
Qy 5461 TGCCAAAGAACTTGAATCCAGGAGCAACGCTAGCTATTTCTGTGCAACCTCTAGAG 5520  
Db 5461 TGCCAAAGAACTTGAATCCAGGAGCAACGCTAGCTATTTCTGTGCAACCTCTAGAG 5520  
Qy 5521 ATTTGCGCTCTCTTCAAAACAAACCCAGGAGGAAAGTAACTAAATCATAAATCCCC 5580  
Db 5521 ATTTGCGCTCTCTTCAAAACAAACCCAGGAGGAAAGTAACTAAATCATAAATCCCC 5580  
Qy 5581 ATGSGCTCTCTTATCATATTTTCTCTKTASTGTTTACCTSTTTTCACTCTCACT 5640  
Db 5581 ATGSGCTCTCTTATCATATTTTCTCTKTASTGTTTACCTSTTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGCAAGCTCCGGAAATATTTGATGCCCATCTGATAGAGTCTTTTAAAGGAAACCCC 5760  
Db 5701 ATGCAAGCTCCGGAAATATTTGATGCCCATCTGATAGAGTCTTTTAAAGGAAACCCC 5760  
Qy 5761 ACCTTCACTGCGCCACACCCATATGCGGCACTGCTATCACTCTGCGCACTTTTGCATG 5820  
Db 5761 ACCTTCACTGCGCCACACCCATATGCGGCACTGCTATCACTCTGCGCACTTTTGCATG 5820  
Qy 5821 CATGCAAAATCACTATTTTGGACAGGAAATGATTAATCTTAGTTCCTGAGGACTT 5880  
Db 5821 CATGCAAAATCACTATTTTGGACAGGAAATGATTAATCTTAGTTCCTGAGGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTGTGAGTGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTGTGAGTGGGGTGGAGTT 5940  
Qy 5941 CAAAGTACAGGCAAGAAAAACATGTAAGAAAGTAATCTCCAACTCAACCSGGGTACAT 6000  
Db 5941 CAAAGTACAGGCAAGAAAAACATGTAAGAAAGTAATCTCCAACTCAACCSGGGTACAT 6000

6001 GGCACTCTAGCCCTCAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACC 6060  
6001 GGCACTCTAGCCCTCAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACC 6060  
6061 CATACTCGCTGGTGAAGCTATTATTAACACACCTCACTGGGCTCCATGAGGTCTCGGCC 6120  
6061 CATACTCGCTGGTGAAGCTATTATTAACACACCTCACTGGGCTCCATGAGGTCTCGGCC 6120  
6121 CAAACCTTACTAATCTTGGATATGCTCCCTCGAATCTTCAAGCCATATGTTCAATC 6180  
6121 CAAACCTTACTAATCTTGGATATGCTCCCTCGAATCTTCAAGCCATATGTTCAATC 6180  
6181 CTGTGCTCTCAACAAATGGAACAACTTCAGCACAGAAATAAACCACTTCCTGTTAGTA 6240  
6181 CTGTGCTCTCAACAAATGGAACAACTTCAGCACAGAAATAAACCACTTCCTGTTAGTA 6240  
6241 GGAACCTCTGTTTCCAAATSTGGAATAAACCACTTCCTGTTAGTA 6300  
6241 GGAACCTCTGTTTCCAAATSTGGAATAAACCACTTCCTGTTAGTA 6300  
6301 AGCAATACTACATACACCAACTCCCAATGATCAGGTGGTAACTCTCCCAACAA 6360  
6301 AGCAATACTACATACACCAACTCCCAATGATCAGGTGGTAACTCTCCCAACAA 6360  
6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTCTGTTG 6420  
6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTCTGTTG 6420  
6421 AATGGCTCTTCAAGATCTATGCTCTCTCAATCTTGTAGTCCCTCCATGTCCTAC 6480  
6421 AATGGCTCTTCAAGATCTATGCTCTCTCAATCTTGTAGTCCCTCCATGTCCTAC 6480  
6481 ACTGAACAGATTTTACAGTTATGTCATATCAATCAAGCCCGCAACAAAGAGTACCATT 6540  
6481 ACTGAACAGATTTTACAGTTATGTCATATCAATCAAGCCCGCAACAAAGAGTACCATT 6540  
6541 CTTCTCTTTGTATAGGACGAGGTCTAGGTGCTAGTCTAGTCTAGTCTAGTCTAGTCT 6600  
6541 CTTCTCTTTGTATAGGACGAGGTCTAGGTGCTAGTCTAGTCTAGTCTAGTCTAGTCT 6600  
6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
6661 GTGCGCGACTCCTGTGCTACCTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCT 6720  
6661 GTGCGCGACTCCTGTGCTACCTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCT 6720  
6721 CBAATCTGAAGACTTTAGACTGCTTAACTGCTGAGAGGGGGAACCTGTTTATTTTA 6780  
6721 CBAATCTGAAGACTTTAGACTGCTTAACTGCTGAGAGGGGGAACCTGTTTATTTTA 6780  
6781 GGGGAAGATGCTGTTATTTATGTTATCAATCGGAATCGTCACTGAGAAAGTTAAGAA 6840  
6781 GGGGAAGATGCTGTTATTTATGTTATCAATCGGAATCGTCACTGAGAAAGTTAAGAA 6840  
6841 ATTCAGATCGAATCAACGTAAGCAGAGAGCTTCGAAACACTGGACCTCGGGCCTC 6900  
6841 ATTCAGATCGAATCAACGTAAGCAGAGAGCTTCGAAACACTGGACCTCGGGCCTC 6900  
6901 CTCAGCCATGATGCTGCTGATCTTCCCTCTTCTAGGACCTCTAGCAGCTATATATG 6960  
6901 CTCAGCCATGATGCTGCTGATCTTCCCTCTTCTAGGACCTCTAGCAGCTATATATG 6960  
6961 CTACTCTCTTTGGACCTGATCTTTTACCTCTGTTAACTTTGCTCTTCCGAAATC 7020  
6961 CTACTCTCTTTGGACCTGATCTTTTACCTCTGTTAACTTTGCTCTTCCGAAATC 7020  
7021 GAAGCTGTAAATCAAAATGGAGCCCAAGATGCAAGTCAAGTCAAGTCAAGTCAAGTCA 7080  
7021 GAAGCTGTAAATCAAAATGGAGCCCAAGATGCAAGTCAAGTCAAGTCAAGTCAAGTCA 7080  
7081 CCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTAATGACATCAAAAGGACCCCTCCT 7140

7081 CCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTAATGACATCAAAAGGACCCCTCCT 7140  
7141 GAGGAAATCTCAGCTGCACAACTCTTACTACGCCCAATTTACGAGGAAACGATGAGC 7200  
7141 GAGGAAATCTCAGCTGCACAACTCTTACTACGCCCAATTTACGAGGAAACGATGAGC 7200  
7201 GGTGTGGGCAACCTCCCAACAGCACTTAGGTCTTCTGTTTGGAGTGGGACTGAGA 7260  
7201 GGTGTGGGCAACCTCCCAACAGCACTTAGGTCTTCTGTTTGGAGTGGGACTGAGA 7260  
7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGATCTTAAGCTTAGTGGGAGGTG 7320  
7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGATCTTAAGCTTAGTGGGAGGTG 7320  
7321 ACCACATCAACCTTTAAACACAGGCTTGCACCTTAGYTACACCTGACCAATCAGAGAG 7380  
7321 ACCACATCAACCTTTAAACACAGGCTTGCACCTTAGYTACACCTGACCAATCAGAGAG 7380  
7381 CTCACATAAATGCTAATAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTC 7440  
7381 CTCACATAAATGCTAATAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTC 7440  
7441 MTGAGAGCAGCAGGAGGACATGATCGGATATAAACCCAAAGTTCGAGCCCGCAA 7500  
7441 MTGAGAGCAGCAGGAGGACATGATCGGATATAAACCCAAAGTTCGAGCCCGCAA 7500  
7501 CGGCAACCCCTTTGGGCTCCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTTAC 7560  
7501 CGGCAACCCCTTTGGGCTCCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTTAC 7560  
7561 TCTATTAAATCTTGCACTGCR 7582  
7561 TCTATTAAATCTTGCACTGCR 7582

RESULT 2  
AA25665  
ID AAX25665 standard; cDNA to mRNA; 7582 BP.  
AC AAX25665;  
XX  
DT 21-MAY-1999 (first entry)  
XX  
Complete human endogenous retrovirus W genome.  
DE  
DE  
XX  
XX  
KW Clone; human endogenous retrovirus; genome; autoimmune disease;  
multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;  
disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.  
XX  
OS Human endogenous retrovirus.  
XX  
XX WO9902696-A1.  
XX  
XX 21-JAN-1999.  
XX  
XX 06-JUL-1998; 98WO-FR001442.  
XX  
XX 07-JUL-1997; 97FR-00008815.  
XX  
XX (INMR ) BIO MERIEUX.  
XX  
XX Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;  
XX  
XX WPI; 1999-120897/10.  
XX  
XX New nucleic acid sequences from human endogenous retrovirus-W - expressed  
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune  
PT disease, and abnormal or failed pregnancy.  
XX  
XX Claim 1; Page 71-74; 106pp; French.  
XX  
XX This sequence represents the complete sequence of the human endogenous

CC retrovirus (HERV) W genome. The nucleic acids, their fragments or  
 CC peptides encoded by them are markers of autoimmune disease (e.g. multiple  
 CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,  
 CC insulin- dependent diabetes and related pathologies) and of abnormal or  
 CC unsuccessful pregnancy and can be used as chromosomal markers for  
 CC susceptibility to these conditions, or proximity markers of genes  
 CC associated with this susceptibility  
 XX

SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;

Query Match 99.9%; Score 7487.4; DB 2; Length 7582;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7578; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy	1	CAACAATCGGAGATATAAACCCAGGCAATTCGAGCTGGCAACGACGACCCCTTTGGGTCC	60
Db	1	CAACAATCGGAGATATAAACCCAGGCAATTCGAGCTGGCAACGACGACCCCTTTGGGTCC	60
Qy	61	CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCACACTGCA	120
Db	61	CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCACACTGCA	120
Qy	121	CTCTCTCGTCCAGTCTTCTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180
Db	121	CTCTCTCGTCCAGTCTTCTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180
Qy	181	TGTTTGGCCACACCCGACACCTGCGCTGACTCCCATCCCTTGGATCCTCGAGGGTGC	240
Db	181	TGTTTGGCCACACCCGACACCTGCGCTGACTCCCATCCCTTGGATCCTCGAGGGTGC	240
Qy	241	CGCTGTGCTCTGATCCAGCGARGCGCCATTGCGCTCCCAATTTGGGTAAAGGCTTGC	300
Db	241	CGCTGTGCTCTGATCCAGCGARGCGCCATTGCGCTCCCAATTTGGGTAAAGGCTTGC	300
Qy	301	CATTGTNCTGCACGGCTAAAGTGCCTGGTGTGTTCTAATTGAGCTGAACACTANTCACT	360
Db	301	CATTGTNCTGCACGGCTAAAGTGCCTGGTGTGTTCTAATTGAGCTGAACACTANTCACT	360
Qy	361	GGGTTCCATGTTCTCTTGTGACCCAGCGGCTCTAATKAACCTATKAACCTTACCACA	420
Db	361	GGGTTCCATGTTCTCTTGTGACCCAGCGGCTCTAATKAACCTATKAACCTTACCACA	420
Qy	421	TGGCCCAAGATTCATCTCTTGGAAATCCGTGAGGSCAACGAACCTCCAGTCAAGAAATAC	480
Db	421	TGGCCCAAGATTCATCTCTTGGAAATCCGTGAGGSCAACGAACCTCCAGTCAAGAAATAC	480
Qy	481	GARGCTTGCCACCATCTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC	540
Db	481	GARGCTTGCCACCATCTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC	540
Qy	541	TTGGGAGCTCTGTGACGACGACCCCGGTGACATTTTGGCRACCAACGAGGACATCC	600
Db	541	TTGGGAGCTCTGTGACGACGACCCCGGTGACATTTTGGCRACCAACGAGGACATCC	600
Qy	601	MAAGTGTGGGAAACGTTCCCGCAAGACAAAAACGCCCTTAAGACGATTTCTGGARAAT	660
Db	601	MAAGTGTGGGAAACGTTCCCGCAAGACAAAAACGCCCTTAAGACGATTTCTGGARAAT	660
Qy	661	TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTGCAGTGCC	720
Db	661	TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTGCAGTGCC	720
Qy	721	GCCTGGCACTCTGTAGGGAAGTATAAATTTAACAACCTTTACAGCTAGACTCTTTTG	780
Db	721	GCCTGGCACTCTGTAGGGAAGTATAAATTTAACAACCTTTACAGCTAGACTCTTTTG	780
Qy	781	TAGAAAAGGCAATGGAGTGAAGTGCATAAGTACAAAACCTTTCTTTTCAATTAAGACAAA	840
Db	781	TAGAAAAGGCAATGGAGTGAAGTGCATAAGTACAAAACCTTTCTTTTCAATTAAGACAAA	840
Qy	841	CTCACAAATTTATGTAAGAGTGTGATTTATGCTTACAGGAAGCCTTCAGAGTCTACCTCC	900
Db	841	CTCACAAATTTATGTAAGAGTGTGATTTATGCTTACAGGAAGCCTTCAGAGTCTACCTCC	900

Qy	901	CTATCCAGCATCCCCGACTCTTCCCAMTYATAAGGACCCCTTCAACCCAAATGG	960
Db	901	CTATCCAGCATCCCCGACTCTTCCCAMTYATAAGGACCCCTTCAACCCAAATGG	960
Qy	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCAAT	1020
Db	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCAAT	1020
Qy	1021	TATGACCCCTCCCAGCAGTGGGAGGAGAAATTCGSCCCAGCCAGAGTGCATGTGCT	1080
Db	1021	TATGACCCCTCCCAGCAGTGGGAGGAGAAATTCGSCCCAGCCAGAGTGCATGTGCT	1080
Qy	1081	TTTTTCTTCCAGACTTTAAAGCAATAAACAAGCACTTAGGTAAATTTCTCAGATAAAYCT	1140
Db	1081	TTTTTCTTCCAGACTTTAAAGCAATAAACAAGCACTTAGGTAAATTTCTCAGATAAAYCT	1140
Qy	1141	GATGGCTATATTGRTGTTTTTACAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAT	1200
Db	1141	GATGGCTATATTGRTGTTTTTACAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAT	1200
Qy	1201	ATATATGCTACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCCACCACTAATGC	1260
Db	1201	ATATATGCTACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCCACCACTAATGC	1260
Qy	1261	AGCCTGAGRGTTTTGGCGATCTCTGATCTCAGTCAGTCAATGGATANGATGACAAACA	1320
Db	1261	AGCCTGAGRGTTTTGGCGATCTCTGATCTCAGTCAGTCAATGGATANGATGACAAACA	1320
Qy	1321	GAAGGAAAGNAATGATTTCCCAACAGCGCAGCAGCAGTCCCACTCTASACCTCATG	1380
Db	1321	GAAGGAAAGNAATGATTTCCCAACAGCGCAGCAGCAGTCCCACTCTASACCTCATG	1380
Qy	1381	GGGACACAGAAATCAGTAAACATGGAGATTTGGTCTGCAGACATTTGCTAATCTGTGTC	1440
Db	1381	GGGACACAGAAATCAGTAAACATGGAGATTTGGTCTGCAGACATTTGCTAATCTGTGTC	1440
Qy	1441	TASAGGACTAAGGAAATCTATGAAATTAATCTCAATGATGTCCACCATTA	1500
Db	1441	TASAGGACTAAGGAAATCTATGAAATTAATCTCAATGATGTCCACCATTA	1500
Qy	1501	ACACAGGGAAGGAAAGAAATCTCTGCTTTTGGAGAGACTAAGGAGGCAATTGAG	1560
Db	1501	ACACAGGGAAGGAAAGAAATCTCTGCTTTTGGAGAGACTAAGGAGGCAATTGAG	1560
Qy	1561	GAAGGCTGCTCTGTGACCTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTT	1620
Db	1561	GAAGGCTGCTCTGTGACCTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTT	1620
Qy	1621	TATCACTCAGTCAGTCGACATTTAGAAAAAACTTCAAAAAAGTCTGCGTAGGCCCGGAG	1680
Db	1621	TATCACTCAGTCAGTCGACATTTAGAAAAAACTTCAAAAAAGTCTGCGTAGGCCCGGAG	1680
Qy	1681	CAAACTTTAGAAAACCTATTGAACTTGGCAACUTCGGTTTTTTTATAATAGAGATCAGGAG	1740
Db	1681	CAAACTTTAGAAAACCTATTGAACTTGGCAACUTCGGTTTTTTTATAATAGAGATCAGGAG	1740
Qy	1741	GAGCAGCGGAACAGGCAAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATCACCT	1800
Db	1741	GAGCAGCGGAACAGGCAAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATCACCT	1800
Qy	1801	CAGGCAAGTGACCTTTGGAGGCTCTGGAAGGAAAAAGCTGGGCAAAATTTGAATGCCATA	1860
Db	1801	CAGGCAAGTGACCTTTGGAGGCTCTGGAAGGAAAAAGCTGGGCAAAATTTGAATGCCATA	1860
Qy	1861	TAGGCTTTGCTTCCAGTCCGCTCTACAAAGGACATTTTAAAAAAGATTTGTCCAAGTAGAG	1920
Db	1861	TAGGCTTTGCTTCCAGTCCGCTCTACAAAGGACATTTTAAAAAAGATTTGTCCAAGTAGAG	1920
Qy	1921	TAGGCGCCCTTCTGCTCATGCCCCCTTATTTCAAGGGAATCACTGGAGGCCCATCTGCC	1980
Db	1921	TAGGCGCCCTTCTGCTCATGCCCCCTTATTTCAAGGGAATCACTGGAGGCCCATCTGCC	1980

1981 QY CAGGGGCAAAAGGTCCTTTGAGTCAGAAAGCCACTTAACCCAGATGATCCAGCAGCAGGACTG 2040  
1981 Db CAGGGGCAAAAGGTCCTTTGAGTCAGAAAGCCACTTAACCCAGATGATCCAGCAGCAGGACTG 2040  
2041 QY AGGGTGCTGGGGCAAGCGCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTT 2100  
2041 Db AGGGTGCTGGGGCAAGCGCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTT 2100  
2101 QY GACCATTTGAGGGCCAGGAAGGTTGTCTCTGGGACACTGGTGGCGGTCTTCTTAGCTTTACT 2160  
2101 Db GACCATTTGAGGGCCAGGAAGGTTGTCTCTGGGACACTGGTGGCGGTCTTCTTAGCTTTACT 2160  
2161 QY CTTCTGTGCCGGCAACACTGTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACG 2220  
2161 Db CTTCTGTGCCGGCAACACTGTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACG 2220  
2221 QY GGCAGTCACCTAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTTT 2280  
2221 Db GGCAGTCACCTAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTTT 2280  
2281 QY TCACATGCTTTTCTTAATTATGCTTGAAGCCCACTACCTTGTGTAGGGAGACATTTCTA 2340  
2281 Db TCACATGCTTTTCTTAATTATGCTTGAAGCCCACTACCTTGTGTAGGGAGACATTTCTA 2340  
2341 QY GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTNCC 2400  
2341 Db GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTNCC 2400  
2401 QY CTTGCTTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAAGGACAAATATGACGAGC 2460  
2401 Db CTTGCTTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAAAGGACAAATATGACGAGC 2460  
2461 QY CAAAGAAATGCCCGTCTGTTCAAGTTAACTAAAGGATTCACCTTCTTCCCTACCAAA 2520  
2461 Db CAAAGAAATGCCCGTCTGTTCAAGTTAACTAAAGGATTCACCTTCTTCCCTACCAAA 2520  
2521 QY GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAGATGTTAAGCACTTAA 2580  
2521 Db GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAGATGTTAAGCACTTAA 2580  
2581 QY AAGCCCAAGGCTTAGTAAACCATGCAATACTCCCTGCACTAATTCGGTAGTGATTGAG 2640  
2581 Db AAGCCCAAGGCTTAGTAAACCATGCAATACTCCCTGCACTAATTCGGTAGTGATTGAG 2640  
2641 QY GAGGCA CAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGATCTCAGGATTAATCAATGG 2700  
2641 Db GAGGCA CAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGATCTCAGGATTAATCAATGG 2700  
2701 QY AGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAAAATAC 2760  
2701 Db AGGCCGTTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAAAATAC 2760  
2761 QY CAGAGGAAGCAGAGTGGTTTACASTCTGACCTTMAGGATGCTTCTTCTGCATCCCTG 2820  
2761 Db CAGAGGAAGCAGAGTGGTTTACASTCTGACCTTMAGGATGCTTCTTCTGCATCCCTG 2820  
2821 QY TACATCCTGACTCTCAATCTTGTGTTGCTTTGAAGATCTTCAAAACCCARCATCTCAAC 2880  
2821 Db TACATCCTGACTCTCAATCTTGTGTTGCTTTGAAGATCTTCAAAACCCARCATCTCAAC 2880  
2881 QY TCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCATCATPATTTGGCCAGGCAT 2940  
2881 Db TCACCTGGACTRTTTTACCCCAAGGGTTCAGGGATAGYCCCATCATPATTTGGCCAGGCAT 2940  
2941 QY TAGCCCAAGACTTGAGYCACTTATACCTGGACACTCTTGTCTTCTGTAGTGGATG 3000  
2941 Db TAGCCCAAGACTTGAGYCACTTATACCTGGACACTCTTGTCTTCTGTAGTGGATG 3000  
3001 QY ATTTACTTTTTRGCGCCYRTTTCAGAAAACCTTGTGCCATCAAGCCACCCAGGCRCTTTTMA 3060  
3001 Db ATTTACTTTTTRGCGCCYRTTTCAGAAAACCTTGTGCCATCAAGCCACCCAGGCRCTTTTMA 3060  
3061 QY ATTTCTCTCGCYACCTGTGGCTACAWGGTTTCCAAACSAARAGCTCARTCTCTGCTCACAGC 3120

3061 Db ATTTCTCTCGCYACCTGTGGCTACAWGGTTTCCAAACSAARAGCTCARTCTCTGCTCACAGC 3120  
3121 QY AGGTTAAATACCTTAGGCRCTAARATTAATCCAAAGCCACCARGGCCCTCAGTGAGGAAYRYA 3180  
3121 Db AGGTTAAATACCTTAGGCRCTAARATTAATCCAAAGCCACCARGGCCCTCAGTGAGGAAYRYA 3180  
3181 QY TCCAGCCTATACCTGGCTTATCTCTCATCYCAAAACCCCTAAGCAACTAAGRRTTCTCTTG 3240  
3181 Db TCCAGCCTATACCTGGCTTATCTCTCATCYCAAAACCCCTAAGCAACTAAGRRTTCTCTTG 3240  
3241 QY GRTTAAYAGGTTCTCCCGAATWATGATTTCCCGAGGTGTGGCRAAATAGCCAGGYCATTA 3300  
3241 Db GRTTAAYAGGTTCTCCCGAATWATGATTTCCCGAGGTGTGGCRAAATAGCCAGGYCATTA 3300  
3301 QY WATACASTAATTAAGGAAACTCAGAAAGCAATACCCATTTARTTAAGATGGAYAMCTGAA 3360  
3301 Db WATACASTAATTAAGGAAACTCAGAAAGCAATACCCATTTARTTAAGATGGAYAMCTGAA 3360  
3361 QY GYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
3361 Db GYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
3421 QY TGCCAAACRGSGCAAGACTTTTSTTTATATYRTACAGAAATAAACAGRAAYAGCTCTRGA 3480  
3421 Db TGCCAAACRGSGCAAGACTTTTSTTTATATYRTACAGAAATAAACAGRAAYAGCTCTRGA 3480  
3481 QY GTCCCTTACACAGRTCCRAGGAYAGCTTGTCAACCCVRTGGRCVACCTGTAAGGAAAT 3540  
3481 Db GTCCCTTACACAGRTCCRAGGAYAGCTTGTCAACCCVRTGGRCVACCTGTAAGGAAAT 3540  
3541 QY GATGTAGTGGCAAGGGTTGRCYTCATTTGTTAYGGGTAGTGTGGCAGTAGCAGTYKTA 3600  
3541 Db GATGTAGTGGCAAGGGTTGRCYTCATTTGTTAYGGGTAGTGTGGCAGTAGCAGTYKTA 3600  
3601 QY GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3601 Db GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3661 QY GTGAAYRGCATACTCACTGTCTAAAGGAGACTTGTGGCTGTGACAGCAACVGTGTACTTAA 3720  
3661 Db GTGAAYRGCATACTCACTGTCTAAAGGAGACTTGTGGCTGTGACAGCAACVGTGTACTTAA 3720  
3721 QY TRTCAAGGCTCTATTACTTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
3721 Db TRTCAAGGCTCTATTACTTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
3781 QY CCAGYCNCAATTTCTTCCAGACAATGAAGAAAGATARAAYATACTGTCAACAARTAAAT 3840  
3781 Db CCAGYCNCAATTTCTTCCAGACAATGAAGAAAGATARAAYATACTGTCAACAARTAAAT 3840  
3841 QY TCTCAAACTATGCCACTCGAGGGGACCTTGTAGAGTTCCYTTGACTGATCCYGNACCTT 3900  
3841 Db TCTCAAACTATGCCACTCGAGGGGACCTTGTAGAGTTCCYTTGACTGATCCYGNACCTT 3900  
3901 QY CAACTTGTATACTGATGGAAGTTCTTTGTAGAAAAGGACTTTCGAAAAGYGGGGTATGC 3960  
3901 Db CAACTTGTATACTGATGGAAGTTCTTTGTAGAAAAGGACTTTCGAAAAGYGGGGTATGC 3960  
3961 QY AGTGTGCTAGTGAATAATGAATTAATTTGAAGTAATCCCTCACTCCAGGAACTAGTGCTYA 4020  
3961 Db AGTGTGCTAGTGAATAATGAATTAATTTGAAGTAATCCCTCACTCCAGGAACTAGTGCTYA 4020  
4021 QY GCTRCAGAACTAATAGCCYTCATYKGGGCACCTAGAAATAGGAGAAAGAAAGGGYAAA 4080  
4021 Db GCTRCAGAACTAATAGCCYTCATYKGGGCACCTAGAAATAGGAGAAAGAAAGGGYAAA 4080  
4081 QY TATATATACAGACTCTRTATATGCTYA CTTAGTCTCATGCTCCCATGMRGCAATATGSA 4140  
4081 Db TATATATACAGACTCTRTATATGCTYA CTTAGTCTCATGCTCCCATGMRGCAATATGSA 4140  
4141 QY AGAAAGGGAATTTCTTAACCTTCYAGRGAACACCTTATCAMAATCAGGAAGCCATTAGGAR 4200



Db 4141 AGAAGGAAATTCCTAACTTCYAGRGAAACACCTATCAMAATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATTATYGGCWGTACAGAAACCTARAGAGTGGWAGTCTTACACTGCGGGGTCACTCA 4260  
Db 4201 ATTATTATYGGCWGTACAGAAACCTARAGAGTGGWAGTCTTACACTGCGGGGTCACTCA 4260  
Qy 4261 NAAAGGAAAGRAAGGGGAAATASAGRGAAATGCGCAAGCAKATATGGAAGWMAAAGAGC 4320  
Db 4261 NAAAGGAAAGRAAGGGGAAATASAGRGAAATGCGCAAGCAKATATGGAAGWMAAAGAGC 4320  
Qy 4321 TGCAAGGAGGAGCCCTCCATTAGAAATGCTTATTAACCTTCCTTTAGTATAGGGTAAATCC 4380  
Db 4321 TGCAAGGAGGAGCCCTCCATTAGAAATGCTTATTAACCTTCCTTTAGTATAGGGTAAATCC 4380  
Qy 4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAAACCTCAGGAGG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAAACCTCAGGAGG 4440  
Qy 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGGAAGGGAATACTTTTTGCCCTGCAAC 4500  
Db 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGGAAGGGAATACTTTTTGCCCTGCAAC 4500  
Qy 4501 TATCAATGGAAATTAATAAACCTTTCATCAAACTTTTCACTTTAGGCATCGATAGCAC 4560  
Db 4501 TATCAATGGAAATTAATAAACCTTTCATCAAACTTTTCACTTTAGGCATCGATAGCAC 4560  
Qy 4561 CCATCARATGGCAATCATATTTACTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGGCAATCATATTTACTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAGGGCTGTGAATGTGCAARAAATAATCCCTGCCCTYATGCGCAAGCTCCTTC 4680  
Db 4621 AKTCAGGGCTGTGAATGTGCAARAAATAATCCCTGCCCTYATGCGCAAGCTCCTTC 4680  
Qy 4681 AGGAAACAAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db 4681 AGGAAACAAARAAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Qy 4741 CCCAAACCTCAGGGATTTTCAGTATCTAATAGTCTGGGTARATACATTTTACCGGGTGGGA 4800  
Db 4741 CCCAAACCTCAGGGATTTTCAGTATCTAATAGTCTGGGTARATACATTTTACCGGGTGGGA 4800  
Qy 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCAAGAGTAAATAAGGCACTAGTTTCATGAA 4860  
Db 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCAAGAGTAAATAAGGCACTAGTTTCATGAA 4860  
Qy 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCCAG 4920  
Db 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCCAG 4920  
Qy 4921 GCCACAGTAAACCCAGGGAGTATCCCAAGGCTTACAGATACGATATCACTTACACTGGGCC 4980  
Db 4921 GCCACAGTAAACCCAGGGAGTATCCCAAGGCTTACAGATACGATATCACTTACACTGGGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCAGGAGGTGCGAATAATGAATGAAYACTCAAGGACATCTA 5040  
Db 4981 TGAAGGCCACAGTCTCAGGAGGTGCGAATAATGAATGAAYACTCAAGGACATCTA 5040  
Qy 5041 AAAAGCAAAACCCAGGAAACCCACCTCACAATGGCCCTGYTCTGTGTGCTATAGCCTTAAAA 5100  
Db 5041 AAAAGCAAAACCCAGGAAACCCACCTCACAATGGCCCTGYTCTGTGTGCTATAGCCTTAAAA 5100  
Qy 5101 AGAATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
Db 5101 AGAATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
Qy 5161 CCTTTCATACCAATGACCTTGTGCTTACCCAGGACCCACTTAGTTTGCAGACATCA 5220  
Db 5161 CCTTTCATACCAATGACCTTGTGCTTACCCAGGACCCACTTAGTTTGCAGACATCA 5220  
Qy 5221 CCTCTTTCAGGCAATATCAACAGTTCTTTAAACATTACAAGGAACCTATCCCTGAGAAG 5280  
Db 5221 CCTCTTTCAGGCAATATCAACAGTTCTTTAAACATTACAAGGAACCTATCCCTGAGAAG 5280

Qy 5281 AGGAAAAGAACTATTCCACCCWMTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATT 5340  
Db 5281 AGGAAAAGAACTATTCCACCCWMTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAITTTATYTTACCCCACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAITTTATYTTACCCCACTG 5400  
Qy 5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGATACATCACTTTGAGTCAAAATCCTGGATAC 5460  
Db 5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGATACATCACTTTGAGTCAAAATCCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACTCTAGAGG 5520  
Db 5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACTCTAGAGG 5520  
Qy 5521 ATTTGCGCTCTCTTCAAAACAACACAGAGGAAAGTAATAAAATCATAAATCCCCC 5580  
Db 5521 ATTTGCGCTCTCTTCAAAACAACACAGAGGAAAGTAATAAAATCATAAATCCCCC 5580  
Qy 5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCTSTTTCACTCTCACT 5640  
Db 5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCTSTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCCCTGTATGACCAAGTAGTCCCTYACCMWAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCCATGCCCTGTATGACCAAGTAGTCCCTYACCMWAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGACGCTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAACCCC 5760  
Db 5701 ATGACGCTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAAACCCC 5760  
Qy 5761 ACCTTCACTGCCACACCCATATGCGCCGAACTGCTATCACTCTGCCACTCTTTGCAATG 5820  
Db 5761 ACCTTCACTGCCACACCCATATGCGCCGAACTGCTATCACTCTGCCACTCTTTGCAATG 5820  
Qy 5821 CATGCAATACATCTATTATGACAGGAAATGATTAATCTCTAGTTGCTTGGAGACTT 5880  
Db 5821 CATGCAATACATCTATTATGACAGGAAATGATTAATCTCTAGTTGCTTGGAGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTATTTACCCAAACTGCTATGCTGTATGGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTATTTACCCAAACTGCTATGCTGTATGGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAAGTAATCTCCCAACTCACCGGGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAAAGTAATCTCCCAACTCACCGGGGTACAT 6000  
Qy 6001 GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 6060  
Db 6001 GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 6060  
Qy 6061 CATACTCCCTGGTAAAGCTATTTAAATACCCCTCACTGGGCTCCATGAGGTCTCGGCC 6120  
Db 6061 CATACTCCCTGGTAAAGCTATTTAAATACCCCTCACTGGGCTCCATGAGGTCTCGGCC 6120  
Qy 6121 CAAAACCTACTAATCTGTGGATATGCTCCCTCTGAACTTCARGCCATATGTTTCAATC 6180  
Db 6121 CAAAACCTACTAATCTGTGGATATGCTCCCTCTGAACTTCARGCCATATGTTTCAATC 6180  
Qy 6181 CCTGTACTGAAACATGGAACAACTTCAGACAGAAATAAACACCACTTCGGTTTTAGTA 6240  
Db 6181 CCTGTACTGAAACATGGAACAACTTCAGACAGAAATAAACACCACTTCGGTTTTAGTA 6240  
Qy 6241 GGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAAAAATTT 6300  
Db 6241 GGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAAAAATTT 6300  
Qy 6301 AGCAATATCATACACAACTCCCAATGCAATGCAATGAGTGGGTAACTCTCTCCACACAA 6360  
Db 6301 AGCAATATCATACACAACTCCCAATGCAATGCAATGAGTGGGTAACTCTCTCCACACAA 6360







Qy	2165	TGTC	CCGGACAACTGTCCTCCAGATCTGTCACTATTCTCAGGGGTTCNNTAAGACGGCA	2224
Db	32501	TGTC	CCGGACAACTGTCCTCCAGATCTGTCACTA-TCTCAGGGGTTC- TAAGACGGCA	32558
Qy	2225	GTCACTAGATAC	TTTTCCTCCAGCCACTAAGTTATGAACTGGGGAGCTTATTCTTTTCAC	2284
Db	32559	GTCACTAGATAC	-TTCTCTCCAGCCACTAAGTTATG- ACTGGGGAGCTTTATTCTTTTCAC	32616
Qy	2285	ATGCTTTTCTAA	TATGCTTTGAAAGCCCACTACCTTGTATTGGGAGAGACATTTCTAGCAA	2344
Db	32617	ATGCTTTTCTAA	TTATGCTTTGAAAGCCCACTACCTTGTATTGGGAGAGACATTTCTAGCAA	32676
Qy	2345	AAGCAGGGCCAT	TATACACCTGAAATAGGAGAGGAACACCCGTTTGTGTTCNCCCTG	2404
Db	32677	AAGCAGGGCCAT	TATACACCTGAAATAGGAGAGGAACACCCGTTTGTGT- CCCTCG	32735
Qy	2405	CTTGAGGAAGAA	TAACTCTGAAGTCTGGCAACAGAGGACAATATGGAGAGCCAAA	2464
Db	32736	CTTGAGGAAGAA	TAACTCTGAAGTCTGGCAACAGAGGACAATATGGAGAG- CAAA	32794
Qy	2465	GAATGCCCGTCT	GTGTTCAAGTTAAACTAAAGGATTCACCTTCCTTCCCTACCAAAGSCA	2524
Db	32795	GAATGCCCGTCT	GTGTTCAAGTTAAACTAAAGGATTCACCTTCCTTCCCTACCAAAGSCA	32854
Qy	2525	GTACCCCTCAG	ACCCCAAGGCCCAACAGGATTCCAAAGATTTGTTAAGGACTTAAAGC	2584
Db	32855	GTACCCCTCAG	ACCCCAAGGCCCAACAGGACTTCCAAAGATTTGTTAAGGACCTAAAGC	32914
Qy	2585	CCAAGGCTTAG	TAAAAACCATGATACCTCCCTGCAGTAATCCGTTAGTGGATTGAGAGG	2644
Db	32915	CCAAGGCTTAG	TAAAAACCATGCAATACCCCTGCAGTACTCC----- AATTTTAGAG	32967
Qy	2645	CACAGAAACCA	GCTGAGCAGTGGAGGGTTAGTGCAGATCTCAGGATTTCAATGGAGGC	2704
Db	32968	TACAGAAACCA	CACAGCAGTGA- GTTTAGTGCAGATCTCAGGATTTCAAT-GAGGC	33025
Qy	2705	CGTTGTCCTTT	ATACCCAGCTGTACTTAGCCCTTATACCTGTGMYTTCCTCCAAATACAGA	2764
Db	33026	TGTTGTTCTCT	ATAGCCAGCTGTACTTAGCCCTTATACCTGTGCTTCCCAATACAGA	33085
Qy	2765	GGAAGCAGAGT	GGTTATACASTCCTCGACCTTGGAGTGCCTTCTCTGCATCCCTGTACA	2824
Db	33086	GGAAGCAGAGT	GGTTTACAGTCTCTGACCTCAGATGCCCTTCTCTGCATCCCTGTACA	33145
Qy	2825	TCCTGACTCTCA	ATTCTTGTTCCTTTGAAAGATACTTCAAACCCARCATCTCAACTCAC	2884
Db	33146	TCCTGACTCTCA	ATTCTTGTTCCTTTGAAAGATACTTCAAACCCARCATCTCAACTCAC	33205
Qy	2885	CTGGACTRTTTT	ACCCAAAGGTTACGGATAGYCCCATCTATTGTCAGGCAATTAGC	2944
Db	33206	CTGGACTRTTTT	ACCCAAAGGTTACGGATAGTCCCCATCTATTGTCAGGCAATTAGC	33265
Qy	2945	CCAAGACTTGAG	YCAPTYMTACACTCGACACTCTTCTCTCTCCTGTAGTGGATGATTT	3004
Db	33266	CCAAGACTTGAG	YCAATCTCATACTCGACA-- CTTGTCTTCGGTAGTGGATGATTT	33323
Qy	3005	ACTTTTTRGVC	GCYRTTCAGAAACCTTGTGCCATCAAGCCACCAAGCRCTCTTTMAATTT	3064
Db	33324	ACTTTTTRGVC	GCYRTTCAGAAACCTTGTGCCATCAAGCCACCAAGCRCTCTTTMAATTT	33383
Qy	3065	CCTCGYACTGT	GTGGTCAAMGGTTTCAAACAPARGCTCACTCTGCTCACAAGCAGGT	3124
Db	33384	CCTCGYACTGT	GTGGTCAAMGGTTTCAAACAPARGCTCACTCTGCTCACAAGCAGGT	33443
Qy	3125	TAAATCTTAGG	BCTHAARATTTACAAAGGCACGAGGCCCTCAGTGGAGAAATATCCA	3184
Db	33444	-----TACTT	AGGGCTAAAATTATCAAAGGCACGAGGCCCTCAGTGGAGAAATATCCA	33499
Qy	3185	GCCTACTGGCT	TATCTCTCYCAAAACCTTAAAGCAATTAAGRBRRTTCTTTGGCT	3244
Db	33500	GCCTACTGGCT	TATCTCTATCCCAAAACCTTAAAGCAATTAAGGGGATTTCTTTGGCT	33559
Qy	3245	AAAYAGGYTT	CTCGCGAANAATGGATTTCCCGAGTTGGCRAAATAGTACGACAGGYCATTAWATA	3304

[illegible]

Db	34623	AGGCAGGACCTCCATTAGAAATGCTTATAAAACAACCCCTAGTATAGGTAAATCCCTC	34682	Qy	5464	CAAAAGAACCTGAAAAATCCAGGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	5523
Qy	4385	CGGGAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGAGG-CAG	4443	Db	35763	CAAAAGAACCTGAAAAATCCAGGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	35822
Db	34683	CGGGAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGAGGACAG	34742	Qy	5524	TGCGCCTGCTCTTCAAAACAACACAGGAGGAAAGTAACATAAAATCATAAATCCCCCATG	5583
Qy	4444	TTTTCTCCCTCGGACGCTTAGCCACTGAAGAAGGAAATACATTTTGGCTTGCACACTAT	4503	Db	35823	TGCGCCTGCTCTTCAAAACAACACAGGAGGAAAGTAACATAAAATCATAAAT-CCCCATG	35881
Db	34743	TTTTCTCCCTCGGACGCTTAGCCACTGAAGAAGGAAATACATTTTGGCTTGCACACTAT	34802	Qy	5584	GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCCTSTTTTCACTCTCATCTGA	5643
Qy	4504	CCAAATGGAAATTAATAAACCTTTCATCAAAACCTTTTCACTTTAGGCATCGATAGCACCA	4563	Db	35882	GCCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCCTSTTTTCACTCTCATCTGA	35941
Db	34803	CCAAATGGAAATTAATAAACCTTTCATCAAAACCTTTTCACTTTAGGCATCGATAGCACCA	34862	Qy	5644	CCCCCTCCATCGCCTGTATGACCAAGTAGCTCCCTTACACAGAGTCTTCTATGGGAATG	5703
Qy	4564	TCARATGGCCAAATCATTTTACTTGGACCAAGGCCCTTTTCAAACTATCAAGCARATKTT	4623	Db	35942	CCCCCTCCATCGCCTGTATGACCAAGTAGCTCCCTTACACAGAGTCTTCTATGGGAATG	36001
Db	34863	TCARATGGCCAAATCATTTTACTTGGACCAAGGCCCTTTTCAAACTATCAAGCARATKTT	34922	Qy	5704	CAGCGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAAACCCCCACC	5763
Qy	4624	CAGGCGCTGTGAATGTGCCARAAATAATCCCTGCTTATCGCCAAAGCTCCCTTCAGG	4683	Db	36002	CAGCGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAAACCCCCACC	36061
Db	34923	CAGGCGCTGTGAATGTGCCARAAATAATCCCTGCTTATCGCCAAAGCTCCCTTCAGG	34982	Qy	5764	TTCACTGCCACACCCCATATATGCCCGCAACTGCTATCACTCTGCGCACTCTTTGCAATGAT	5823
Qy	4684	APAAACAAARAAACAGGCCATTACCTTGRARAAACTGGCAACTGATTTTACCCACAAAGCCC	4743	Db	36062	TTCACTGCCACACCCCATATATGCCCGCAACTGCTATCACTCTGCGCACTCTTTGCAATGAT	36121
Db	34983	AGAACAAAGAACAGGCCATTACCTTGGAGAGACTGGCAACTGATTTTACCCACAAAGCCC	35042	Qy	5824	GCAAATCTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTTGA	5883
Qy	4744	AAACCTCAGGGATTTTCACTATCTAGTCTGGGTATATATCTTCAACGGGTTGGCABAG	4803	Db	36122	GCAAATCTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTTGA	36181
Db	35043	AAACCTCAGGGATTTTCACTATCTAGTCTGGGTATATATCTTCAACGGGTTGGCABAG	35102	Qy	5884	GTCACTGCTCTGTGGACTTTACTTCAACCAACTGTATGTCTGATGGGGTGGAGTTCAA	5943
Qy	4804	GCTTTCCCTGTAGACAGAAAGCCCAAGAGTGAATAAAGGCACTAGTTTCAATGAATA	4863	Db	36182	GTCACTGCTCTGTGGACTTTACTTCAACCAACTGTATGTCTGATGGGGTGGAGTTCAA	36241
Db	35103	GCTTTCCCTGTAGACAGAAAGCCCAAGAGTGAATAAAGGCACTAGTTTCAATGAATA	35162	Qy	5944	GATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACATGCG	6003
Qy	4864	ATTCCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTTCCAGGCC	4923	Db	36242	GATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTACATGCG	36301
Db	35163	ATTCCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTTCCAGGCC	35222	Qy	6004	ACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCAT	6063
Qy	4924	ACAGTAACCCAGGGAGTATCCAGGGGTTAGTATACGNATATCACTTACACTGCGCCTGA	4983	Db	36302	ACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCAT	36361
Db	35223	ACAGTAACCCAGGGAGTATCCAGGGGTTAGTATACGNATATCACTTACACTGCGCCTGA	35282	Qy	6064	ACTGCGCTGTGAAGCTATTTTAATACCACTCCTCCTGCGGCTCCATGAGTCTCGGCCCAA	6123
Qy	4984	AGGCCACAGTCTCTAGGAGAGTTCAGAAATGAATGAATGAATGAATGAATGAATGAAT	5043	Db	36362	ACTGCGCTGTGAAGCTATTTTAATACCACTCCTCCTGCGGCTCCATGAGTCTCGGCCCAA	36421
Db	35283	AGGCCACAGTCTCTAGGAGAGTTCAGAAATGAATGAATGAATGAATGAATGAATGAAT	35342	Qy	6124	AACCCCTACTAATGTGGATATGCTCCCTGAACTTTTCAAGCCATATGTTTCAATCCCT	6183
Qy	5044	AAGCAAAACCCAGGAAACCCACTCAGTGGCCTGTCTGTGCTTATAGCCCTTAAAGAGA	5103	Db	36422	AACCCCTACTAATGTGGATATGCTCCCTGAACTTTTCAAGCCATATGTTTCAATCCCT	36481
Db	35343	AAGCAAAACCCAGGAAACCCACTCAGTGGCCTGTCTGTGCTTATAGCCCTTAAAGAGA	35402	Qy	6184	GTACCTGAACTGGAACAACTTCAGCAGAGAAATAAACACCACTTCGTTTACTAGGA	6243
Qy	5104	ATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	5163	Db	36482	GTACCTGAACTGGAACAACTTCAGCAGAGAAATAAACACCACTTCGTTTACTAGGA	36541
Db	35403	ATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	35462	Qy	6244	CCTCTGTTTCCCAATSTGGAATTAACCCATACCTCAAACTCACCTGTGTAAAAATTTAGC	6303
Qy	5164	TTCAATACCAATGACTTGTGCTTGAACCAAGACCAACTTAGTTGGAGACATCACT	5223	Db	36542	CCTCTGTTTCCCAATSTGGAATTAACCCATACCTCAAACTCACCTGTGTAAAAATTTAGC	36601
Db	35463	TTCAATACCAATGACTTGTGCTTGAACCAAGACCAACTTAGTTGGAGACATCACT	35522	Qy	6304	AATACTACATACACAAACCACTCCCAATGCAATGAGGTGGGTAACTCTCTCCCAACAAATA	6363
Qy	5224	CCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAGAACTTATCCCTGAGAAGAGG	5283	Db	36602	AATACTACATACACAAACCACTCCCAATGCAATGCAATGAGGTGGGTAACTCTCTCCCAACAAATA	36661
Db	35523	CCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAGAACTTATCCCTGAGAAGAGG	35582	Qy	6364	GTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCCTATCGTTTGAAT	6423
Qy	5284	GAAGAAGAACTATTCACCCWGTGACATGGTATTAAGTCAAGTCCCTTCTTAATTC	5343	Db	36662	GTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCCTATCGTTTGAAT	36721
Db	35583	GAAGAAGAACTATTCACCCWGTGACATGGTATTAAGTCAAGTCCCTTCTTAATTC	35642	Qy	6424	GGCTCTTCAGAACTATGTGCTTCTCTCATTCTTAGTGCCCTCATGRCCATCTACACT	6483
Qy	5344	CATCCCTAGATACATCTGGAGGACCCCTACCCAGTCAATTTATTTATCCCAACTGCGG	5403	Db	36722	GGCTCTTCAGAACTATGTGCTTCTCTCATTCTTAGTGCCCTCATGRCCATCTACACT	36781
Db	35643	CATCCCTAGATACATCTGGAGGACCCCTACCCAGTCAATTTATTTATCTACCCCACTGCGG	35702	Qy	6484	GAACAAGATTTATACAGTTATGTCTATATCTAAGCCCGCCCAACAAAGAGTACCCATCTTT	6543
Qy	5404	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACAATCTTGAATCAAAATCCTGATCTGC	5463	Db	36782	GAACAAGATTTATACAGTTATGTCTATATCTAAGCCCGCCCAACAAAGAGTACCCATCTTT	36841
Db	35703	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACAATCTTGAATCAAAATCCTGATCTGC	35762				

```
QY 6544 CCTTTTGTATAGGAGCAGGAGTCTAGGTGCACTAGTACTGCGATTGGCGGTATCACA 6603
Db 36842 CCTTTTGTATAGGAGCAGGAGTCTAGGTGCACTAGTACTGCGATTGGCGGTATCACA 36901
QY 6604 ACCTCTACTCAGTTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC 6663
Db 36902 ACCTCTACTCAGTTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC 36961
QY 6664 GCCGACTCCCTGGTGCACCTTGCAGAGTCAACTTAACTCCCTAGCAGAGTAGTCCCTTCA 6723
Db 36962 GCCGACTCCCTGGTGCACCTTGCAGAGTCAACTTAACTCCCTAGCAGAGTAGTCCCTTCA 37021
QY 6724 AATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGGGGGAACCTGTTTATTTTAGGG 6783
Db 37022 AATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGGGGGAACCTGTTTATTTTAGGG 37081
QY 6784 GAAGAATGCTGTTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843
Db 37082 GAAGAATGCTGTTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 37141
QY 6844 CSAGATCGAATACAACTAGCAGAGAGGAGCTTCGAAACACTGAGCCCTGGGGCCCTCCTC 6903
Db 37142 CGAGATCGAATACAACTAGCAGAGAGGAGCTTCGAAACACTGAGCCCTGGGGCCCTCCTC 37201
QY 6904 AGCCRAATGATGCCCTGGATTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTTGCTA 6963
Db 37202 AGCCRAATGATGCCCTGGATTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTTGCTA 37261
QY 6964 CTCCTCTTTGGACCTGTATCTTTTACCTCTTTTAACTTTGTTCTCTCCAGAAATCGAA 7023
Db 37262 CTCCTCTTTGGACCTGTATCTTTTAACTCTTTGTTTAACTTTGTTCTCTCCAGAAATCGAA 37321
QY 7024 GCTGTAAACTACAAATGAGGCCCAAGAGTGCAGTCCAAAGACTAAGATCTACCCGACCC 7083
Db 37322 GCTGTAAACTACAAATGAGGCCCAAGAGTGCAGTCCAAAGACTAAGATCTACCCGACCC 37381
QY 7084 CTGGACCGGCTGTTAGCCCACTCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAG 7143
Db 37382 CTGGACCGGCTGTTAGCCCACTCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAG 37441
QY 7144 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTTACGACGAGGAGTGTAGAGCGGT 7203
Db 37442 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTTACGACGAGGAGTGTAGAGCGGT 37501
QY 7204 SGTGGCCCACTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGACTGAGAGAC 7263
Db 37502 CGTGGCCCACTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGACTGAGAGAC 37561
QY 7264 AGGACTAGCTGGATTTCTAGGCTGATTAAGATCCCTTAAGCTAGTGGGAGGTGACC 7323
Db 37562 AGGACTAGCTGGATTTCTAGGCTGATTAAGATCCCTTAAGCTAGTGGGAGGTGACC 37621
QY 7324 ACATCCACCTTTTAAACACGGGGCTTCAACTTAGTCAACCTGACCAATTCAGAGAGCTC 7383
Db 37622 ACATCCACCTTTTAAACACGGGGCTTCAACTTAGTCAACCTGACCAATTCAGAGAGCTC 37681
QY 7384 ACTAAATGCTAATTAGGCAAGACAGAGGAGTAAAGAAATAGCCATCATYTTATTCGCTG 7443
Db 37682 ACTAAATGCTAATTAGGCAAGACAGAGGAGTAAAGAAATAGCCATCATYTTATTCGCTG 37741
QY 7444 AGAGCAGCAGAGGAGGACATGATCGGATATAAACCAGATGTTTCGAGCCGGCAACGG 7503
Db 37742 AGAGCAGCAGAGGAGGACATGATCGGATATAAACCAGATGTTTCGAGCCGGCAACGG 37801
QY 7504 CAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
Db 37802 CAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCACTCT 37861
QY 7564 ATTAAATCTTGCARCTGCR 7582
Db 37862 ATTAAATCTTGCARCTGCA 37880
```

```
RESULT 4
ABN97929
ID ABN97929 standard; DNA; 10499 BP.
XX
AC ABN97929;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence HERV-7q.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
multiple sclerosis; ds.
XX
OS Human endogenous retrovirus.
XX
PN WO967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR001513.
XX
PR 23-JUN-1998; 98FR-00007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
WP1; 2000-160587/14.
XX
New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
for diagnosis, treatment and prevention of autoimmune and neurological
diseases.
Claim 3; Fig 1; 225pp; French.
The present invention relates to new nucleic acid sequences of human
endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
Regulatory elements associated with HERV-7q may alter expression of other
genes (even remote genes) on the same chromosome, inducing immunological
and/or neurological changes (which may be pathological or protective/
curative). HERV-7q peptides can be used to improve efficiency of the
immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
sequences can be used in immunogenic or vaccinating compositions, for
protection against autoimmune diseases, particularly multiple sclerosis.
The peptides may also be used (by sequence comparison) to detect/identify
endogenous retroviruses that are abnormally expressed in cancer,
neuropathologies or other autoimmune diseases. The present sequence was
used to illustrate the invention
SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 U; 0 Other;
```

```
Query Match 85.9%; Score 6436; DB 3; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

QY 606 GATGGAAACGTTCCCGCAAGCAAAAACGCCCTTAAGACGTATTCTCGAATAATGGGA 665
Db 2953 GATGGAAACGTTCCCGCAAGCAAAAACGCCCTTAAGACGTATTCTCGAATAATGGGA 3012
QY 666 MCAATTTGACCTCGACACTAAGAAAGAAAGCACTTATTCTTCTGCAAGTCCGCCCTG 725
Db 3013 CCAATTTGACCTCGACACTAAGAAAGAAAGCACTTATTCTTCTGCAAGTCCGCCCTG 3072
QY 726 GCACCTCCTGAGGAGTAGTAATTAACACCATCTTACAGCTAGACCTCTTTTGTAGAA 785
Db 3073 GCACCTCCTGAGGAGTAGTAATTAACACCATCTTACAGCTAGACCTCTTTTGTAGAA 3132
QY 786 AAGGCAATGGAGTGAAGTGCATTAAGTACAACTTTCTTTTTCATTAAAGACAACTCAC 845
Db 3133 AAGGCAATGGAGTGAAGTGCATTAAGTACAACTTTCTTTTTCATTAAAGACAACTCAC 3192
QY 846 AATTATGTAAGTGTGATTATCCCTTACAGGAGCCTTCAGAGTCTACCTCCCTATC 905
```





3065 QY CCTCGVACCTGTGGCTACAWGGTTTCCAAACCSABARGCTCARTCTGTCTCAGCAGGT 3124  
5384 Db CCTCGTACTCTGTGGCTACATGTTTCCAAACCAAGGCTCAACTCTGTCTCAGCAGGT 5443  
3125 QY TAAATCTTAGGCTAARATTTATCCAAAGGCACCAAGGCCCTCAGTCAGGAAAYRATCCA 3184  
5444 Db ----TACTTAGGCTAATAATTATCCAAAGGCACCAAGGCCCTCAGTCAGGAAACACATCCA 5499  
3185 QY GCCTATACCTGCTTATCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 3244  
5500 Db GCCTATACCTGCTTATCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 5559  
3245 QY AAYAGGVTTCGCGCAAWATGGATTCCCGAGGTTCGCGCAAWATGGATTTCGCGCAAWATGGATT 3304  
5560 Db AATAGGTTTCGCGCAAWATGGATTCCCGAGGTTCGCGCAAWATGGATTTCGCGCAAWATGGATT 5618  
3305 QY CASTAATTAAGGAACTCAGAAAGCCCAATACCAATTTARTAGATGGAYAMCTGAAGYMR 3364  
5619 Db CACTAATTAAGGAACTCAGAAAGCCCAATACCAATTTARTAGATGGAYAMCTGAAGYMR 5678  
3365 QY AAGTGGCTTTCAGGCCCCCTAAGAGGCTTAAACCCCAAGYCCCAAGTGTGAAGYTTGCC 3424  
5679 Db AAGTGGCTTTCAGGCCCCCTAAGAGGCTTAAACCCCAAGYCCCAAGTGTGAAGYTTGCC 5724  
3425 QY AACRGGGCAAGACTTTTSTTATATVATTCACAGAAACCAAGCAAGCAAGCAAGCAAGCAAGCA 3484  
5725 Db AACRGGGCAAGACTTTTSTTATATVATTCACAGAAACCAAGCAAGCAAGCAAGCAAGCAAGCA 5783  
3485 QY TTACACAGATCCGAGGGATGAGCTTGCACCTGTGGCATACCTGACTAAGGAAATTTGATG 3544  
5784 Db TTACACAGATCCGAGGGATGAGCTTGCACCTGTGGCATACCTGACTAAGGAAATTTGATG 5843  
3545 QY TAGTGGCAAGGTTGRCVTCATPTGTTTAYVGGTGTGTGGCAGTGTGAGTGTGATGAT 3604  
5844 Db TAGTGGCAAGGTTGRCVTCATPTGTTTAYVGGTGTGTGGCAGTGTGAGTGTGATGAT 5903  
3605 QY CTGAGCACTTAAATAATACAGGAGAGATCTTACTGTGGCATCTCATGAGTGA 3664  
5904 Db CTGAGCACTTAAATAATACAGGAGAGATCTTACTGTGGCATCTCATGAGTGA 5963  
3665 QY AYRGCACTACTGCTGAAGGAGACTTGTGGCTGTGCACCACTGTGTTTACTTAAATRTC 3724  
5964 Db AYRGCACTACTGCTGAAGGAGACTTGTGGCTGTGCACCACTGTGTTTACTTAAATRTC 6023  
3725 QY AGGCTCTATTACTTGAAGGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3784  
6024 Db AGGCTCTATTACTTGAAGGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6083  
3785 QY YCNCACTTCTCAGACATAGAAAGATARAATATATATATATATATATATATATATATATATAT 3844  
6084 Db YCNCACTTCTCAGACATAGAAAGATARAATATATATATATATATATATATATATATATATAT 6143  
3845 QY AAACCTATCCACTCGAGGGGACCTTATAGAGTTCCTTATAGAGTTCCTTATAGAGTTCCTTAT 3904  
6144 Db AAACCTATCCACTCGAGGGGACCTTATAGAGTTCCTTATAGAGTTCCTTATAGAGTTCCTTAT 6202  
3905 QY TTGTATCTAGTGGAGTTCTTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964  
6203 Db TTGTATCTAGTGGAGTTCTTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 6262  
3965 QY GTCAGTGAATAGGAATAATTTGAAGTAATCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 4024  
6263 Db GTCAGTGAATAGGAATAATTTGAAGTAATCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 6322  
4025 QY GCAGAACTAATGCTTATGCTTACCTGTGCTTACCTGTGCTTACCTGTGCTTACCTGTGCTTAC 4084  
6323 Db GCAGAACTAATGCTTATGCTTACCTGTGCTTACCTGTGCTTACCTGTGCTTACCTGTGCTTAC 6382  
4085 QY TATACAGACTTATATATATGCTTACCTGTGCTTACCTGTGCTTACCTGTGCTTACCTGTGCT 4144  
6383 Db TATACAGACTTATATATATGCTTACCTGTGCTTACCTGTGCTTACCTGTGCTTACCTGTGCT 6442  
4145 QY AGGGAATTCCTAACTTCYAGRGAACACCTATCAMAACATCAGGAAGCCATTAGGARATTA 4204

6443 Db AGGGAATTCCTAACTTCYAGRGAACACCTATCMAACATCAGGAAGCCATTAGGAATTA 6502  
4205 QY TTATYGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCTGCGGGTCTATCANAAA 4264  
6503 Db TTATYGGCWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCTGCGGGTCTATCANAAA 6562  
4265 QY GGAAAGRAAGGAAATTAAGAAATTCGCAAGCAKATATTTGAAGCMAAAAGAGCTGCA 4324  
6563 Db GGAAAGRAAGGAAATTAAGAAATTCGCAAGCAKATATTTGAAGCMAAAAGAGCTGCA 6622  
4325 QY AGGCAAGGACCTCCCTATAGAAATGCTTATTAACCTTCCCTTAGTATAGGTAATCCCTTC 4384  
6623 Db AGGCAAGGACCTCCCTATAGAAATGCTTATTAACCTTCCCTTAGTATAGGTAATCCCTTC 6682  
4385 QY CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGG - CAG 4443  
6683 Db CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGGAG - CAG 6742  
4444 QY TTTTCTCCCTCGGACGGTTAGCCACTGGAAGAGGAAATACTTTTGGCTGCAACTAT 4503  
6743 Db TTTTCTCCCTCGGACGGTTAGCCACTGGAAGAGGAAATACTTTTGGCTGCAACTAT 6802  
4504 QY CCAATGGGAAATTTACTTAAACCCCTTCATCAAAACCTTTTCACTTTAGGCATCGATAGCAG 4563  
6803 Db CCAATGGGAAATTTACTTAAACCCCTTCATCAAAACCTTTTCACTTTAGGCATCGATAGCAG 6862  
4564 QY TCARATGGCCAAATCACTTATTTACTTGGACAGGCTTTTCAAACTATCAAGCAATAT 4623  
6863 Db TCARATGGCCAAATCACTTATTTACTTGGACAGGCTTTTCAAACTATCAAGCAATAT 6922  
4624 QY CAGGCTCTGTGAATGTGCAARAAATAATCCCTGCTATCGCAAGCTCTTTCAGG 4683  
6923 Db CAGGCTCTGTGAATGTGCAARAAATAATCCCTGCTATCGCAAGCTCTTTCAGG 6982  
4684 QY ABAAACAAABACAGGCCATTTACCTGTRARAPACTTGGCACTGATTTTACCAAGGCC 4743  
6983 Db ABAAACAAABACAGGCCATTTACCTGTRARAPACTTGGCACTGATTTTACCAAGGCC 7042  
4744 QY AAACCTCAGGGATTTTCACTATCTAGTCTGGGTARATCTTTCACGGTGGGCGABAG 4803  
7043 Db AAACCTCAGGGATTTTCACTATCTAGTCTGGGTARATCTTTCACGGTGGGCGABAG 7102  
4804 QY GCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATAATA 4863  
7103 Db GCCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATAATA 7162  
4864 QY ATTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGAATATAGCCCTTTCAGGCC 4923  
7163 Db ATTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGAATATAGCCCTTTCAGGCC 7222  
4924 QY ACAGTAACCCAGGGAGTATCCAGGCTTACGATATACGATATACCTTACCTGCGCTGA 4983  
7223 Db ACAGTAACCCAGGGAGTATCCAGGCTTACGATATACGATATACCTTACCTGCGCTGA 7282  
4984 QY AGGCAACAGTCTCTCAGGGAAGGTCCAGAAAATGAATGAATAAAVACTCAAGGACATCTAAA 5043  
7283 Db AGGCAACAGTCTCTCAGGGAAGGTCCAGAAAATGAATGAATAAAVACTCAAGGACATCTAAA 7342  
5044 QY AAGCAAAACCCAGGAAACCCACTCAGTGGCTGTCTGTTGCTTATAGCCCTTAAAAAGA 5103  
7343 Db AAGCAAAACCCAGGAAACCCACTCAGTGGCTGTCTGTTGCTTATAGCCCTTAAAAAGA 7402  
5104 QY ATCTGCACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGGCC 5163  
7403 Db ATCTGCACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGGCC 7462  
5164 QY TTCTAAACCAATGACTTGTGCTTGAACCAAGCAAGCAACTTACTGTGAGACATCACT 5223  
7463 Db TTCTAAACCAATGACTTGTGCTTGAACCAAGCAAGCAACTTACTGTGAGACATCACT 7522  
5224 QY CTTTAGCCAAATATCAACAGTTCCTTAAAAACATTAAGGAAGCCCTTATCCCTGAGAGAGG 5283



Db 7523 CCTAGCCAAATATCAACAAGTTCTTAAAAAATTAACAAGGAACCTATCCCTGAGAAGAGG 7582  
Qy 5284 GAAAGAAGACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTTCCC 5343  
Db 7583 GAAAGAAGACTATTCACCCCTTTGTGACATGGTATTTAGTCAAGTCCCTTCCTCTAATTTCCC 7642  
Qy 5344 CATCCCTAGATACATCTCCGGAAGGACCTACCCAGTCAATTTATTTATACCCCAACTGGGG 5403  
Db 7643 CATCCCTAGATACATCTCCGGAAGGACCTACCCAGTCAATTTATTTATCTACCCCAACTGGGG 7702  
Qy 5404 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATCTGC 5463  
Db 7703 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGGATCTGC 7762  
Qy 5464 CAAAGGAACCTGAAATCTCAGAGACAAAGCTAGTATTTCTGTGAACTCTAGAGGAT 5523  
Db 7763 CAAAGGAACCTGAAATCTCAGAGACAAAGCTAGTATTTCTGTGAACTCTAGAGGAT 7822  
Qy 5524 TGCCTCTCTCTTCAAAACAAACAGGAGGAAGTAATAAAATCATAAATCCCCCATG 5583  
Db 7823 TGCCTCTCTCTTCAAAACAAACAGGAGGAAGTAATAAAATCATAAATCCCCCATG 7881  
Qy 5584 GSCCTCCCTATCATATTTCTCTKATSTGTSTTTTACCTCTTTCACCTCTCACTGCA 5643  
Db 7882 GSCCTCCCTATCATATTTCTCTTCTTACTTGTCTTTTACCTCTTTCACCTCTCACTGCA 7941  
Qy 5644 CCCCCTCCATGCGCTGTATGACACAGTAGTCCCTCTYACWAGAGTTTCTATGGAGATG 5703  
Db 7942 CCCCCTCCATGCGCTGTATGACACAGTAGTCCCTCTTACCAAGAGTTTCTATGGAGATG 8001  
Qy 5704 CAGCTGCCGGAATATTTGATGCCCATCGTATAGAGTCTTTSTAAGGGAACCCCCACC 5763  
Db 8002 CAGCTGCCGGAATATTTGATGCCCATCGTATAGAGTCTTTCTAAGGGAACCCCCACC 8061  
Qy 5764 TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAT 5823  
Db 8062 TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAT 8121  
Qy 5824 GCAATATCATTTATGGACAGGAAAAATGATTAATCTCTAGTTCTCTGGAGGACTTGG 5883  
Db 8122 GCAATATCATTTATGGACAGGAAAAATGATTAATCTCTAGTTCTCTGGAGGACTTGG 8181  
Qy 5884 GTCACTGTCTGTGGACTTACTTCCACCAACTGCTATGCTCTGATGGGGGTGGAGTTCAA 5943  
Db 8182 GTCACTGTCTGTGGACTTACTTCCACCAACTGCTATGCTCTGATGGGGGTGGAGTTCAA 8241  
Qy 5944 GATCAGGCAAGAAAAAATGATAAAGAGTAATCTCCCAACTCACCGGGGTACATGGC 6003  
Db 8242 GATCAGGCAAGAAAAAATGATAAAGAGTAATCTCCCAACTCACCGGGGTACATGGC 8301  
Qy 6004 ACCTCTAGCCCTACAAGGACTAGATCTCTCAAAATACATGAAACCTCTCGTTACCCAT 6063  
Db 8302 ACCTCTAGCCCTACAAGGACTAGATCTCTCAAAATACATGAAACCTCTCGTTACCCAT 8361  
Qy 6064 ACTCCCTGGTAAGCTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCCCAA 6123  
Db 8362 ACTCCCTGGTAAGCTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCCCAA 8421  
Qy 6124 AACCTCTACTAGTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 6183  
Db 8422 AACCTCTACTAGTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 8481  
Qy 6184 GTACTGAAACAAATGGAACAACTTCAGCACAGAAAAATAAACACCACTTCCTGTTTAGTGA 6243  
Db 8482 GTACTGAAACAAATGGAACAACTTCAGCACAGAAAAATAAACACCACTTCCTGTTTAGTGA 8541  
Qy 6244 CCTCTGTGTTTCCAAATSTGGAATACCCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
Db 8542 CCTCTGTGTTTCCAAATSTGGAATACCCATACCTCAAACTCACCTGTGTAAATTTAGC 8601  
Qy 6304 AATACTACATACAAACCAACTCCCAATGCATCAGTGGGTAACTCTCCCAACACAAATA 6363  
Db 8602 AATACTACATACAAACCAACTCCCAATGCATCAGTGGGTAACTCTCCCAACACAAATA 8661

Qy 6364 GTCTGCTACCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGAAT 6423  
Db 8662 GTCTGCTACCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGAAT 8721  
Qy 6424 GGCTCTTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCCYATGRCATCTACACT 6483  
Db 8722 GGCTCTTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCCCTATGACCATCTACACT 8781  
Qy 6484 GAAACAATTTATACAGTTATGTCAATCTTAAGCCCCCGCAACAAAAGAGTACCAATCTTT 6543  
Db 8782 GAAACAATTTATACAGTTATGTCAATCTTAAGCCCCCGCAACAAAAGAGTACCAATCTTT 8841  
Qy 6544 CCTTTCTGTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCACA 6603  
Db 8842 CCTTTCTGTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCACA 8901  
Qy 6604 ACCTCTACTCAGTTTCTACTACAAATCTCTCAAGAACTAAATGGGGACATCGAAGCGGTC 6663  
Db 8902 ACCTCTACTCAGTTTCTACTACAAATCTCTCAAGAACTAAATGGGGACATCGAAGCGGTC 8961  
Qy 6664 GCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCAA 6723  
Db 8962 GCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCAA 9021  
Qy 6724 AATCGAAGAGCTTTAGACTTGTCTAACCCTGABAGGGGGAACCTCTTTATTTTAGGG 6783  
Db 9022 AATCGAAGAGCTTTAGACTTGTCTAACCCTGABAGGGGGAACCTCTTTATTTTAGGG 9081  
Qy 6784 GAAGAATGCTGTATTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 9082 GAAGAATGCTGTATTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 9141  
Qy 6844 CSAGATCGAATPACAACTAKAGCAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCTCT 6903  
Db 9142 CGAGATCGAATPACAACTAGAGCAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCTCT 9201  
Qy 6904 AGCCBATGATGCTCCCTGGATTCCTCCCTTCTTAGAGCTCTAGCAGCTATATATTTGCTA 6963  
Db 9202 AGCCBATGATGCTCCCTGGATTCCTCCCTTCTTAGAGCTCTAGCAGCTATATATTTGCTA 9261  
Qy 6964 CTCTCTTTGGACCTCTATCTTTTACCTCTCTGTAACTTTGTCTCTTCAGAAATCGAA 7023  
Db 9262 CTCTCTTTGGACCTCTATCTTTTAACTCTCTGTAACTTTGTCTCTTCAGAAATCGAA 9321  
Qy 7024 GCTGTAAAATPACAAATGAGGCCCAAGATGCAAGTCCAAGACTAAGATCTACCGCAGACC 7083  
Db 9322 GCTGTAAAATPACAAATGAGGCCCAAGATGCAAGTCCAAGACTAAGATCTACCGCAGACC 9381  
Qy 7084 CTGGAACCGGCTGVTAGCCCACTGATCTGATGTTAATGACATCAAGGCAACCCCTCCTGAG 7143  
Db 9382 CTGGAACCGGCTGVTAGCCCACTGATCTGATGTTAATGACATCAAGGCAACCCCTCCTGAG 9441  
Qy 7144 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCCGT 7203  
Db 9442 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCCGT 9501  
Qy 7204 SGTCCGCCCACTCCCAACACGACTTAGGTTTCTGTTGAGATGGGGGCTAGAGAAC 7263  
Db 9502 C-TGGGCCCACTCCCAACACGACTTAGGTTTCTGTTGAGATGGGGGCTAGAGAAC 9560  
Qy 7264 AGGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCCTAGGCTAGTGGGAAGGTGACC 7323  
Db 9561 AGGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCCTAGGCTAGTGGGAAGGTGACC 9620  
Qy 7324 ACATCCACTTTAAACACGGGCTTTGCAACTTAGTGTACACTGACCAATCAGAGAGCTC 7383  
Db 9621 ACATCCACTTTAAACACGGGCTTTGCAACTTAGTGTACACTGACCAATCAGAGAGCTC 9680  
Qy 7384 ACTAAAATGCTTAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATTTATTTGCTG 7443  
Db 9681 ACTAAAATGCTTAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATCTATTGCTG 9740

QY 7444 AGACACAGCAGGAGGACATGATCGGATATAAACCCAGTTCGAGCCGCAACGG 7503  
 Db |||||||  
 QY 9741 AGACACAGCAGGAGGACATGATCGGATATAAACCCAGTTCGAGCCGCAACGG 9800  
 Db |||||||  
 QY 7504 CAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 7563  
 Db |||||||  
 QY 9801 CAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 9860  
 QY 7564 ATTAATCTTGCARCTGCR 7582  
 Db |||||||  
 QY 9861 ATTAATCTTGCARCTGCA 9879  
 Db |||||||

RESULT 5  
 ID ADS30988 standard; DNA; 8523 BP.  
 XX  
 AC ADS30988;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human genome high complexity repeat found in the HIRA gene #21.  
 XX  
 KW Human; ds;  
 KW histone cell cycle regulation defective, *S. cerevisiae* homologue A; HIRA;  
 KW high complexity repeat; in situ hybridisation; Southern blot;  
 KW chromosome breakpoint; inherited genetic disease; neoplastic disorder;  
 KW chromosome 22; DiGeorge syndrome; Velo-Cardio-facial syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003224356-A1.  
 XX  
 PD 04-DEC-2003.  
 XX  
 PF 14-MAY-2001; 2001US-00854867.  
 XX  
 PR 16-MAY-2000; 2000US-00573080.  
 XX  
 PA (KNOLL/) KNOLL J H M.  
 PA (ROGA/) ROGAN P K.  
 PI Knoll JHM, Rogan PK;  
 XX  
 DR WPI; 2002-062378/08.  
 XX  
 PT Single copy genomic hybridization probes for detecting specific nucleic  
 PT acid sequences in sample by in situ hybridization useful for detection of  
 PT acquired or inherited genetic diseases.  
 XX  
 PS Example 1; SEQ ID NO 21; 30pp; English.  
 XX  
 CC The invention relates to a nucleic acid hybridisation probe comprising a  
 CC labelled, single copy nucleic acids of at least 50 nucleotides, which  
 CC will hybridise to a deduced single copy sequence interval in target  
 CC nucleic acid (TNA) of known sequence. The single copy sequence is deduced  
 CC by comparing the target nucleic acid (e.g. a disease causing gene) with a  
 CC collection of high and low complexity repeat sequences as found in the  
 CC genome of the organism from containing the target nucleic acid. The probe  
 CC is generated by PCR on the target sequence. The probe is essentially free  
 CC of blocking nucleic acid sequences which will hybridise to repeat  
 CC sequences within the genome of which the TNA is a part, and is labelled  
 CC with a label selected from fluorochrome-responsive labels, fluorochromes,  
 CC calorimetric chemical, conjugated proteins, antibodies, antigens and  
 CC their mixtures. The probe is useful in a hybridisation method, where the  
 CC hybridisation method is from in situ hybridisation, Southern blot, and  
 CC other methods in which nucleic acid is immobilised, where the method  
 CC further comprises selecting a single copy nucleic acid which will  
 CC hybridise to a duplicate or triplicon sequence domain. The probe is useful  
 CC for determining the existence of previously unknown repeat sequence  
 CC families in a genome. The method comprises reacting a labelled probe with  
 CC the genome, causing the probe to hybridise and ascertaining if the probe  
 CC hybridises to the genome at more than three preferably ten different

CC locations as a determination of new repeat sequence family, where the  
 CC determining step comprises selecting the single copy sequence from a  
 CC duplicon or triplicon sequence domain. The probe is useful for  
 CC determining a chromosome breakpoint and is useful in the fields for  
 CC cytogenetics and molecular genetics for determining the presence of  
 CC specific nucleic acid sequences in a sample of eukaryotic origin, e.g.  
 CC the probes may be used to analyse specific chromosomal locations by in  
 CC situ hybridisation as a detection of acquired or inherited genetic  
 CC diseases especially for detection of genetic or neoplastic disorders.  
 CC Unlike prior art techniques, the probe permits more precise chromosomal  
 CC breakpoint determinations by in situ hybridisation. The genomic sequence  
 CC comprising the human HIRA gene (histone cell cycle regulation defective,  
 CC *S. cerevisiae*, homologue A) was analysed for single copy sequence  
 CC intervals for use as probes of the invention. HIRA is located on  
 CC chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and  
 CC Velo-Cardio-facial syndromes. The present sequence is a high complexity  
 CC repeat found within the human genome used to analyse the HIRA gene for  
 CC repeat regions. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030224356.  
 XX  
 SQ Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;

Query Match 81.8%; Score 6124.6; DB 7; Length 8523;  
 Best Local Similarity 95.7%; Pred. No. 0;  
 Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGGAACGTTTCCCGCAAGACAAACCGCCCTTAAGACGTTATTCGGAATTTGGGA 665  
 Db |||||||  
 QY 1923 GATGGGAACGTTTCCCGCAAGACAAACCGCCCTTAAGACGTTATTCGGAATTTGGGA 1982  
 Db |||||||  
 QY 666 MCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCGAGTCCGCGCTG 725  
 Db |||||||  
 QY 1983 CCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCGAGTCCGCGCTG 2042  
 Db |||||||  
 QY 726 GCATCTCTGAGGAGTATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAA 785  
 Db |||||||  
 QY 2043 GCATCTCTGAGGAGTATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAA 2102  
 Db |||||||  
 QY 786 AAGGCAATGGAGTGAAGTGCCATAGTAAGTCAAACTTTCTTTTCATTAAGAGACAACCTCAC 845  
 Db |||||||  
 QY 2103 AAGGCAATGGAGTGAAGTGCCATAGTAAGTCAAACTTTCTTTTCATTAAGAGACAACCTCAC 2162  
 Db |||||||  
 QY 846 AATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC 905  
 Db |||||||  
 QY 2163 AATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC 2222  
 Db |||||||  
 QY 906 CCAGCATCCCGACTCTTCCCGCAAGTAAATAGGAGCCCGCTTCAACCCAAATGGTCCAA 965  
 Db |||||||  
 QY 2223 CCAGCATCCCGACTCTTCCCGCAAGTAAATAGGAGCCCGCTTCAACCCAAATGGTCCAA 2282  
 Db |||||||  
 QY 966 AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAATTTATGA 1025  
 Db |||||||  
 QY 2283 AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAATTTATGA 2342  
 Db |||||||  
 QY 1026 CCCCTCCCAAGCAGTGGGAGAGAGAAATTCGCGCCAGCAGAGTGCATGTGCTTTT 1085  
 Db |||||||  
 QY 2343 CCCCT-CCAAGCAGTGGGAGAGAGAAATTCGCGCCAGCAGAGTGCATGTGCTTTTTC 2401  
 Db |||||||  
 QY 1086 TCTCCAGACTTAAAGCAAAATAAACAAGACTTAGGTAATTTCTCAGATAAATCTCATGG 1145  
 Db |||||||  
 QY 2402 TCTCCAGACTTAAAGCAAAATAAACAAGACTTAGGTAATTTCTCAGATAAATCTCATGG 2461  
 Db |||||||  
 QY 1146 CTATATCTGTGTTTACAGGGTGTAGGCAATTTCTTGATCTGACATGAGAGATATATA 1205  
 Db |||||||  
 QY 2462 CTATATCTGTGTTTACAGGGTGTAGGCAATTTCTTGATCTGACATGAGAGATATATA-A 2520  
 Db |||||||  
 QY 1206 TGTCTACTGCTAAATCAGACACTAAACCCCAATAGAGAGAGTGCCACATAATCGAGCCT 1265  
 Db |||||||  
 QY 2521 TGTCTACTGCTAAATCAGACACTAAACCCCAATAGAGAGAGTGCCACATAATCGAGCCT 2580  
 Db |||||||  
 QY 1266 GAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGACCAACGAAG 1325  
 Db |||||||

Db 2581 GAGAGTTTGGGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAAACAG-AGG 2637  
Qy 1326 AAAGANAATGATTTCCCAAGCCAGCAGCAGCGAGTTCACAGTCTASACCTCATTTGGGGAC 1385  
Db 2638 AAAGAGAATGATTTCCCAAGCCAGCAGCAGCGAGTTCACAGTCTAGACCTCATTTGGG--- 2694  
Qy 1386 ACAGAAATCAGTAACATGGAGATGGTGTCTGCAGACATTTGCTAACTTGTGTGTCTASAA 1445  
Db 2695 ACACAGATCAGAACAT-GGAGATTTGGTGTCTGCAGACATTTGCTAACTTGTGTGTCTAGNA 2753  
Qy 1446 GGACTAAGGAAAACCTASGAAGAAATCTAYGAATTTACTCAATGATGTCCACCATAACACA 1505  
Db 2754 GGACTAAGGAAAACCTTAGGAAG-AGTCTATGAATTTACTCAATGATGTCCACCATAACACA 2812  
Qy 1506 GGGGAAGGAGAGAAAATCTTACTGCTTCTGAGAGAGACTAAGGAGGCAATTGAGGAAGC 1565  
Db 2813 -GGGAAGGAGAGAAAATCTTACTGCTTCTGAGAGAGACTAAGGAGGCAATTGAGGAAGC 2871  
Qy 1566 GTGCCCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 1625  
Db 2872 GTGCCCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 2931  
Qy 1626 CTCAGTCAGCTGCAGACATTAG-AAAAAATTTCAAAGTCTGCCGTAGGCCCGGAGCAAA 1684  
Db 2932 CTCAGTCAGCTGCAGACATTAGAAAAAATTTCAAAGTCTGCCGTAGGCCCGGAGCAAA 2991  
Qy 1685 ACTTAGAAACCTTATGAACTTGGCAACVTCGGTTTTTTTATATAGATCAGGAGGAGC 1744  
Db 2992 ACTTAGAAACCTTATGAACTTGGCAACCTCGTTTTTTTATATAGATCAGGAGGAGC 3051  
Qy 1745 AGGCGGAACAGGACAAACGGGATTAAGAAAGGCCACCGCTTTAGTCATGACCCCTCAGG 1804  
Db 3052 AGGCGGAACAGGACAAACGGGATTAAGAAAGGCCACCGCTTTAGTCATGACCCCTCAGG 3111  
Qy 1805 CAAAGTGGACTTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAATTGAATGCTAATAGG 1864  
Db 3112 CAAAGTGGACTTTGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAATTGAATGCTAATAGG 3171  
Qy 1865 GTTGTCTTCCAGTGGGCTTCAAGGACACTTTTAAAAAGATGTGCCAAGTAGAAGTAAG 1924  
Db 3172 GTTGTCTTCCAGTGGGCTTCAAGGACACTTTTAAAAAGATGTGCCAAGTAGAAGTAAG 3231  
Qy 1925 CGGCCCTTCGTCATGCCCTTATTAAGGGATCACTGGAAGGCCCACTGCCCCAGG 1984  
Db 3232 CGGCCCTTCGTCATGCCCTTATTAAGGGATCACTGGAAGGCCCACTGCCCCAGG 3291  
Qy 1985 GGACAAAGTCTTTTGAAGTCAGAGCCTAACACAGATGATCCAGCAGCAGGACTGAGGG 2044  
Db 3292 GGACAAAGTCTTCTGATCAGAGGCCACTAACAGATGATCCAGCAGCAGGACTGAGGG 3351  
Qy 2045 TGCTGGGGCAAGGCCCATCCATGCCATCACCTTCACAGAGCCTTGGGTATGCTTGACC 2104  
Db 3352 TGCTGGGGCAAGGCCCATCCATGCCATCACCTTCACAGAGCCTTGGGTATGCTTGACC 3411  
Qy 2105 ATTGAGGGCCAGAGGTTGTCTCTGGACACTGGTGGGCTTCTTACTTACTCTTC 2164  
Db 3412 ATTGAGGGCCAGG-AGGTTGTCTCTGGACACTGGTGGGCTTCTTACTTACTCTTC 3470  
Qy 2165 TGTCGGGCAACACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGGCA 2224  
Db 3471 TGTCGGGCAACACTGTCTCCAGATCTGTCACTA-TCTGAGGGGTCC-TAAGACGGGCA 3528  
Qy 2225 GTCACTAGATCTTTTCCAGCCACTAAGTTTATGAATGAGCTGGGAGCTTTATCTTTTCAC 2284  
Db 3529 GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTGGGGAGCTTTATCTTTTCAC 3586  
Qy 2285 ATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTAGCNA 2344  
Db 3587 ATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTAGCNA 3646  
Qy 2345 AAGCAGGGGCCATTAACACCTGAAACATAGGAGAAACACCCCGTTGTTGTCNCCCTG 2404  
Db 3647 AAGCAGGGGCCATTAACACCTGAAACATAGGAGAAACACCCCGTTGTTGTCNCCCTG 3705

Qy 2405 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGACAATATGGACGAGCCAAA 2464  
Db 3706 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGACAATATGGACGAG-CAAA 3764  
Qy 2465 GAATGCCCGTCTCTCTCAAGTTTAAACTAAAGGATTTCCATCTTCTTCCCTTACCRAAGCA 2524  
Db 3765 GAATGCCCGTCTCTCTCAAGTTTAAACTAAAGGATTTCCACCTCTTCTTCCCTTACCRAAGCA 3824  
Qy 2525 GTACCCCTCAGACCCCAAGGCCCAACAGAGGATTTCAAAAAGATTTGTTAAGGACTTAAAGC 2584  
Db 3825 GTACCCCTCAGACCCCAAGGCCCAACAGGACTCTCAAAAAGATTTGTTAAGGACTTAAAGC 3884  
Qy 2585 CCAAGGCTTAGTAAAAACATGCATAACTCCCTGAGTAATTCCTGAGTGGATTTGAGGAGG 2644  
Db 3885 CCAAGGCTTAGTAAAAACATGCAGTAACCCCTCGAGTACTCC-----AATTTTAGGAG 3937  
Qy 2645 CACAGAAACCCAGTGGGACAGTGGAGGGTTAGTGCAGAGATCTCAGGATTTATCAATGGAGGC 2704  
Db 3938 TACAGAAACCAACAGACAGTGGAG-GGTTAGTGCAGAGATCTCAGGATTTATCAAT-GAGGC 3995  
Qy 2705 CGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTTCCCAAAATACAGA 2764  
Db 3996 TGTGTTCTCTATAGCCAGCTGTACTAGCCCTTATATCTGTCTTCTCCAAATACAGA 4055  
Qy 2765 GGAAGCAGAGTGTTTTACASTCCTGGACCTTMAGGATGCTTCTCTGCATCCCTGTACA 2824  
Db 4056 GGAAGCAGAGTGTTTTACAGTCTCGGACCTTCAGAGTCCCTTCTCTGCATCCCTGTACA 4115  
Qy 2825 TCCTGACTCTCAATCTTGTGCTTTTGAAGATACTTTCAAACCCARCATCTCAACTCAC 2884  
Db 4116 TCCTGACTCTCAATCTTGTGCTTTTGAAGATACTTTCAAACCCCAACATCTCAACTCAC 4175  
Qy 2885 CTGACATTTTATACCCAAAGGTTTCAAGGATGAGYCCCATCTATTTGGCCAGGCAATTAGC 2944  
Db 4176 CTGACTATTTTATACCCAAAGGTTTCAAGGATGAGTCCCATCTATTTGGCCAGGCAATTAGC 4235  
Qy 2945 CCAAGACTTGAGYCARVTWTCATACCTGGACACTCTTGTCTCTCTGTAAGTGGATGATTT 3004  
Db 4236 CCAAGACTTGAGCAATCTCTCATCTCTGGAACA--CTTGTCTCTGTAAGTGGATGATTT 4293  
Qy 3005 ACTTTTRGVCYCCYRVTTCAGAAAACCTTGTGCATCAAGCCCAAGCCCTCTTMAATTT 3064  
Db 4294 ACTTTTGGCCGCCATTCAGAAAACCTTGTGCATCAAGCCCAAGCCCTCTTCAATTT 4353  
Qy 3065 CCTCGACACTGTGGCTACAWGGTTTCAAAACSPARGCTCACTCTCTCTCACAGAGT 3124  
Db 4354 CCTCGACTGTGGCTACATGTTTCCAAACCAAGGCTCAACTCTCTCTCACAGAGT 4413  
Qy 3125 TAATACTTAGGCTTAATATTTATCCAAAGGCACCAAGGCCCTCAGTGAAGGAAATATCCA 3184  
Db 4414 ----TACTTAGGGCTTAAATTTATCCAAAGGCACCAAGGCCCTCAGTGAAGGAAATATCCA 4469  
Qy 3185 GCCTATCTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRGRRRTTCTTTGGGCT 3244  
Db 4470 GCCTATCTGGCTTATCTCTCATCCCAAAACCTTAAAGCAACTAAGGGGATTTCTTTGGCT 4529  
Qy 3245 AAYAGGYTTTGGCCGAATGATTTCCCGATGTTGCGRAAATAGCCAGGYCATTAATA 3304  
Db 4530 AATAGGTTTCTGCCGAAATGGAAT-CCAGGATGCGGAAATAGCCAGGTCATTAATA 4588  
Qy 3305 CASTAATTAAGGAAACTCAGAAAGCCATCCATTTTARTAGATGGAATGAGYACTGAAGYMR 3364  
Db 4589 CACTAATTAAGGAAACTCAGAAAGCCATCCATTTTATAGTGAAGTGAACACTGAAGTAG 4648  
Qy 3365 AAGTGGCTTTCCAGGCC-----TAACCCAGCCCAAGTGAAGTTGCTTGGC 3424  
Db 4649 AAGTGGCTTTCCAGGCC-----TAACCCAGCCCAAGTGAAGTTGCTTGGC 4694  
Qy 3425 AACRGGGCAAGACTTTTAT 3484  
Db 4695 AACRGGGCAAGACTTTTCTCATATGTATGTATGTATGTATGTATGTATGTATGTATGTATGT 4753

3485 TTACAGRCCCRAGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGCAAAATGATG 3544  
4754 TTACACAGATCCCGAGGATGAGCTTGCAACCTCTGGCCATACCTGACTAAGGAAATGATG 4813  
3545 TAGTGGCAAAAGGGTTGRCYCTCATTTTAYGGGTAGTGGGAGTAGCAGTGYKTAGTAT 3604  
4814 TAGTGGCAAAAGGGTTGACCTCATTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT 4873  
3605 CTGAGCAGTTAAATAATACAGGAGAGATCTTACTGTGTGACATCTCATGAKGTGA 3664  
4874 CTGAGCAGTTAAATAATACAGGAGAGATCTTACTGTGTGACATCTCATGATGTGA 4933  
3665 AYRGCTACTCACTGTCTAAAGGAGACTTGTGGCTGTGCAGCAACYGTGTTACTTAAATRTC 3724  
4934 ATGGCATCTCACTGTCTAAAGGAGACTTGTGGCTGTGCAGCAACYGTGTTACTTAAATGTC 4993  
3725 AGGCTCTATTACTGTAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAACCCAG 3784  
4994 AGGCTCTATTACTGTAARGGCCAGTGTGCGACTGTGCACTTGTGCAACTCTTAAACCCAG 5053  
3785 YCNCAATTTCTCAGACAATGAAGAAAGATARAATAATCTGTCAACAARTATTTCTC 3844  
5054 CCACATTTCTCAGACAATGAAGAAAGATARAATAATCTGTCAACAARTATTTCTC 5113  
3845 AAACCTATGCACTCGAGGGGACCTTGTAGAGTTCCYTTGACTGTATCCYVGCCTTCAAC 3904  
5114 AAACCTATGCACTCGAGGGGACCTTGTAGAGTTCCYTTGACTGTATCCYVGCCTTCAAC 5172  
3905 TTGTATCTGATGGAAGTTCTCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964  
5173 TTGTATCTGATGGAAGTTCTCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 5232  
3965 GTCAGTGATATGGAATATYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTR 4024  
5233 GTCAGTGATATGGAATATYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTR 5292  
4025 GCAGAACTAATACCCCTCACTTGGGCACCTAGAAATAGGAGAGAAAGGGYAAATATA 4084  
5293 GCAGAACTAATACCCCTCACTTGGGCACCTAGAAATAGGAGAGAAAGGGYAAATATA 5352  
4085 TATACAGACTCTAARTATGCTYACCTAGTCTNTCCATGCCCATGMRCAATATGSRAGAA 4144  
5353 TATACAGACTCTAARTATGCTYACCTAGTCTNTCCATGCCCATGMRCAATATGSRAGAA 5412  
4145 AGGGAATTCCTAATCTCYGAGGAAACCTATCAMACATCAGGAAGCCATAGGABATTA 4204  
5413 AGGGAATTCCTAATCTCYGAGGAAACCTATCAMACATCAGGAAGCCATAGGABATTA 5472  
4205 TTATYTCGCHGTACAGAACTTACAGAGGTGGAGTCTTACACTGCGGGTCAATCAAAA 4264  
5473 TTATYTCGCHGTACAGAACTTACAGAGGTGGAGTCTTACACTGCGGGTCAATCAAAA 5532  
4265 GGAAAGAAAGGAAATASAGRGAAVTGCAAGCAKATATTGAAGCMAAAGAGCTGCA 4324  
5533 GGAAAGAAAGGAAATAGAGAACTGCAAGCAGATATTGAAGCMAAAGAGCTGCA 5592  
4325 AGCAGGACCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC 4384  
5593 AGCAGGACCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC 5652  
4385 CGGGAACCCAGCCAGTACTCAGCAGGAGAAACAGATGGGGAACCTCAGAGG - CAG 4443  
5653 CGGGAACCCAGCCAGTACTCAGCAGGAGAAACAGATGGGGAACCTCAGAGGACAG 5712  
4444 TTTTCTCCCTCGGGAGGTTAGCCACTGAGAGGGGAAATATCTTTTCCCTGCACTAT 4503  
5713 TTTTCTCCCTCGGGAGGTTAGCCACTGAGAGGGGAAATATCTTTTCCCTGCACTAT 5772  
4504 CCAATGGAATTTACTTTAAACCCCTTCAATCAAACTTTCACTTAGGCATCGATAGCACCA 4563  
5773 CCAATGGAATTTACTTTAAACCCCTTCAATCAAACTTTCACTTAGGCATCGATAGCACCA 5832  
4564 TCARATGGCCAAATCATTTATTTACTGGAACGAGCCCTTTTCAAACTATCAAGCARATAT 4623

5833 TCAGATGGCCAAATCATTTATTTACTGGACGAGCCCTTTTCAAACTATCAAGCAGATAGT 5892  
4624 CAGGGCTGTGAATCTGCGCARAAAAATAATCCCTGCTCTYATCGCCAGAGCTCTTCAGG 4683  
5893 CAGGGCTGTGAATCTGCGCARAAAAATAATCCCTGCTCTYATCGCCAGAGCTCTTCAGG 5952  
4684 ARACAAARAAACAGGCCATTACCTGGRARAACTCGGCAACTGATTTTACCCACAAGCCC 4743  
5953 AGAACAAAGAACAGGCCATTACCTGGRARAACTCGGCAACTGATTTTACCCACAAGCCC 6012  
4744 AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTCAGGGTTGGGAGAG 4803  
6013 AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTCAGGGTTGGGAGAG 6072  
4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA 4863  
6073 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA 6132  
4864 ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGA CAATAGCCCTGCTTTCCAGGCC 4923  
6133 ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGA CAATAGCCCTGCTTTCCAGGCC 6192  
4924 ACAGTAACCCAGGAGTATCCCGAGGCTTACAGAGTGA CAATAGCCCTGCTTTCCAGGCC 4983  
6193 ACAGTAACCCAGGAGTATCCCGAGGCTTACAGAGTGA CAATAGCCCTGCTTTCCAGGCC 6252  
4984 AGGCCACAGTCTCTCAGGGAAGGTTCAGAAAAATGAATGAAYACTCAAAGGCATCTTAAAA 5043  
6253 AGGCCACAGTCTCTCAGGGAAGGTTCAGAAAAATGAATGAAYACTCAAAGGCATCTTAAAA 6312  
5044 AAGCAAAACCCAGGAAACCCACCTCACTGCGCTGTCTGCTGCTATAGCCTTAAAAAGA 5103  
6313 AAGCAAAACCCAGGAAACCCACCTCACTGCGCTGTCTGCTGCTATAGCCTTAAAAAGA 6372  
5104 ATCTGCACTTTTCCCAAAAGCAGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 5163  
6373 ATCTGCACTTTTCCCAAAAGCAGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 6432  
5164 TTCTAATCAATGACCTGTGTGTGACCAAGACAGCACTTGTGTGAGACATCACCT 5223  
6433 TTCTAATCAATGACCTGTGTGTGACCAAGACAGCACTTGTGTGAGACATCACCT 6492  
5224 CTTTAGCCAAATATCAACAAAGTTCTTAAACATTAACAGGAACCTATCCCTGAGAGAGG 5283  
6493 CTTTAGCCAAATATCAACAAAGTTCTTAAACATTAACAGGAACCTATCCCTGAGAGAGG 6552  
5284 GAAAGAACTATTCCACCCWGTGACATGTTATAGTCAAGTCCCTTCTCTTAATCCC 5343  
6553 GAAAGAACTATTCCACCCWGTGACATGTTATAGTCAAGTCCCTTCTCTTAATCCC 6612  
5344 CATCCCTAGATACATCCCTGGGAAGGACCTTACCCAGTCAATTTATCTACCCCACTGCGG 5403  
6613 CATCCCTAGATACATCCCTGGGAAGGACCTTACCCAGTCAATTTATCTACCCCACTGCGG 6672  
5404 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGAAGTCAATCTCTGATACTGC 5463  
6673 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGAAGTCAATCTCTGATACTGC 6732  
5464 CAAAGAACTGTAATAATCCAGGAGCAACGCTAGTATTTCTGTGAACCTCTAGAGATT 5523  
6733 CAAAGAACTGTAATAATCCAGGAGCAACGCTAGTATTTCTGTGAACCTCTAGAGATT 6792  
5524 TCGGCTCTCTTCAACCAACAGGAGGAGTAACTAAATCATAAATCCCCCATG 5583  
6793 TCGGCTCTCTTCAACCAACAGGAGGAGTAACTAAATCATAAATCCCCCATG 6851  
5584 GSCCTCCCTTATCATATTTTCTKTATGTTTCTTACCTCTTCTCTCTCTCTCTCTCTGCA 5643  
6852 GSCCTCCCTTATCATATTTTCTKTATGTTTCTTACCTCTTCTCTCTCTCTCTCTCTGCA 6911  
5644 CCCCTCCATGCGCTGTATGACAGTACTCCCTTACCMAGAGTTTCTATGGAATG 5703

Db 6912 CCCCCTCCATGCCGCTGTATGACCACTAGCTCCCTTACCAAGAGTTTCTTATGGAGATG 6971  
Qy 5704 CAGCCTCCGGAAATATTGATGCCCCATCGTATAGAGTCTTTTAAAGGAACCCCCACC 5763  
Db 6972 CAGCGTCCGGAAATATTGATGCCCCATCGTATAGAGTCTTTTAAAGGAACCCCCACC 7031  
Qy 5764 TTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACATCTTTGATGTCAT 5823  
Db 7032 TTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACATCTTTGATGTCAT 7091  
Qy 5824 GCAAACTACTCAATATTGGACAGAGAAAATGATTAATCTCTAGTGTCTCTGGAGGACTTGG 5883  
Db 7092 GCAAACTACTCAATATTGGACAGAGAAAATGATTAATCTCTAGTGTCTCTGGAGGACTTGG 7151  
Qy 5884 GTCACTGTCTGTGGACTTACTTCAACCAAACTGTGTATGCTGTAGTGGGGGTGGAGTTCAA 5943  
Db 7152 GTCACTGTCTGTGGACTTACTTCAACCAAACTGTGTATGCTGTAGTGGGGGTGGAGTTCAA 7211  
Qy 5944 GATCAGGCAAGAGAAAACATGTAAGAAGTAATCTCCCAACTCACCGSGGTACATGCC 6003  
Db 7212 GATCAGGCAAGAGAAAACATGTAAGAAGTAATCTCCCAACTCACCGSGGTACATGCC 7271  
Qy 6004 ACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 6063  
Db 7272 ACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 7331  
Qy 6064 ACTCGCTGTGAAGCTATTAAATACCACTCTCACTGGGCTCCATGAGGTCTCGGCCCAA 6123  
Db 7332 ACTCGCTGTGAAGCTATTAAATACCACTCTCACTGGGCTCCATGAGGTCTCGGCCCAA 7391  
Qy 6124 RACCCCTACTACTGTGTGATATGCCCTCCCTGAACTTCARGCCATATGTTCAATCCCT 6183  
Db 7392 RACCCCTACTACTGTGTGATATGCCCTCCCTGAACTTCARGCCATATGTTCAATCCCT 7451  
Qy 6184 GTACCTGAACCAATGGAACAACTTCAGCACAGAAATAAAACACACCTTCCTGGTTTATAGGA 6243  
Db 7452 GTACCTGAACCAATGGAACAACTTCAGCACAGAAATAAAACACACCTTCCTGGTTTATAGGA 7511  
Qy 6244 CTTCTTTGTTTCAATSTGGAATAACCCATACCTCAACCTCACTGTTGTAATAATTTAGC 6303  
Db 7512 CTTCTTTGTTTCAATSTGGAATAACCCATACCTCAACCTCACTGTTGTAATAATTTAGC 7571  
Qy 6304 RATACTACATACACACCAACTCCCAATGCATCAGGTGGTAACCTCCACACAAATA 6363  
Db 7572 RATACTACATACACACCAACTCCCAATGCATCAGGTGGTAACCTCCACACAAATA 7631  
Qy 6364 GTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCCCTATCGTTGTTGAAT 6423  
Db 7632 GTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCCCTATCGTTGTTGAAT 7691  
Qy 6424 GGCTCTTCAGAAATCTATGTCTCTCTCATTTTGTAGTGCCCCVATGRCATCTACACT 6483  
Db 7692 GGCTCTTCAGAAATCTATGTCTCTCTCATTTTGTAGTGCCCCVATGRCATCTACACT 7751  
Qy 6484 GNAAGAGATTATACAGTTTGTCTATCTAAGCCCGCAACAAAGATACCCATTTCT 6543  
Db 7752 GNAAGAGATTATACAGTTTGTCTATCTAAGCCCGCAACAAAGATACCCATTTCT 7811  
Qy 6544 CTTTTTGTATAGGAGCAGGAGTGTAGGTGACCTAGGTACTGTGGCATTTGGCGGTATACA 6603  
Db 7812 CTTTTTGTATAGGAGCAGGAGTGTAGGTGACCTAGGTACTGTGGCATTTGGCGGTATACA 7871  
Qy 6604 ACCTCTACTCAGTTCTACTACAAATCTATCTCAAGAACTAAATGGGGAATGGAACGGGTC 6663  
Db 7872 ACCTCTACTCAGTTCTACTACAAATCTATCTCAAGAACTAAATGGGGAATGGAACGGGTC 7931  
Qy 6664 GCGGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCRA 6723  
Db 7932 GCGGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCRA 7991  
Qy 6724 AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGGAACTGTTTATTTTATAGG 6783  
Db 7992 AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGGAACTGTTTATTTTATAGG 8051

Qy 6784 GAAGAACTGCTTTATTTATTTATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAATT 6843  
Db 8052 GAAGAACTGCTTTATTTATTTATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAATT 8111  
Qy 6844 CSAGATCGAATACAACTGAKAGCAGAGAGCTTCGAAACACTGGACCTCGGGCTCTCTC 6903  
Db 8112 CGAGATCGAATACAACTGAGCAGAGAGCTTCGAAACACTGGACCTCGGGCTCTCTC 8171  
Qy 6904 AGCCATATGATGCTCTGATTTCTCCCTTTCTTAGGACCTCTTAGCAGCTTATATTTGCTA 6963  
Db 8172 AGCCATATGATGCTCTGATTTCTCCCTTTCTTAGGACCTCTTAGCAGCTTATATTTGCTA 8231  
Qy 6964 CTCTCTTTGGACCTGTATCTTTTACCTCTCTGTTTAACTTTGTTCTCTTCAGAAATCGAA 7023  
Db 8232 CTCTCTTTGGACCTGTATCTTTTAACTCTCTGTTTAACTTTGTTCTCTTCAGAAATCGAA 8291  
Qy 7024 GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTTAAGATCTACCGCAGACC 7083  
Db 8292 GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTTAAGATCTACCGCAGACC 8351  
Qy 7084 CTGGACCGGCTGYTAGCCCACTCTGATGTTTAACTTAATGACATCAAAAGGCACCCCTCTCTGAG 7143  
Db 8352 CTGGACCGGCTGYTAGCCCACTCTGATGTTTAACTTAATGACATCAAAAGGCACCCCTCTCTGAG 8411  
Qy 7144 GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203  
Db 8412 GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 8471  
Qy 7204 SGTGGCCCACTCCCAACAGCACTTAGTGTTCCTGTTGAGATGGGGAC 7255  
Db 8472 GGTGGCCCACTCCCAACAGCACTTAGTGTTCCTGTTGAGATGGGGAC 8523

RESULT 6  
ID ADY36376 standard; DNA; 8523 BP.  
XX  
AC ADY36376;  
XX  
DT 05-MAY-2005 (first entry)  
XX  
DE HIRA genomic fragment SEQ ID NO 21.  
XX hybridization; DNA detection; neoplasm; genetic disorder; cytogenetics;  
KW HIRA; ds.  
XX Homo sapiens.  
XX  
XX WO200188089-A2.  
XX  
XX 22-NOV-2001.  
XX  
XX 15-MAY-2001; 2001WO-US015674.  
XX  
PR 16-MAY-2000; 2000US-00573080.  
PR 14-MAY-2001; 2001US-00854867.  
XX  
XX (CHIL-) CHILDREN'S MERCY HOSPITAL.  
XX  
XX Knoll JHM, Rogan PK, Cazarro PM;  
XX  
XX WPI; 2002-062378/08.  
XX  
PT Single copy genomic hybridization probes for detecting specific nucleic  
PT acid sequences in sample by in situ hybridization useful for detection of  
PT acquired or inherited genetic diseases.  
XX  
XX Example 1; SEQ ID NO 21; 67pp; English.  
XX  
CC The invention describes a nucleic acid hybridization probe (I) comprising  
CC a labeled, single copy nucleic acids of at least 50 nucleotides, which  
CC will hybridize to a deduced single copy sequence interval in target

2462 CTATATTGATGTTTTACAGGGTTAGACAATCTTTGATCTGCATGGAGAGATATA- A 2520

1206 TGTCTACTGCTAAATCAGACACTAAACCCCAAAATGAGAGAAGTGCCACCATAACTCAGCGCT 1265

2521 TGTCACTGCTAAATCAGACACTAAACCCCAAAATGAGAGAAGTGCCACCATAACTCAGCGCT 2580

1266 GAGRGTTGGCGATCTCTGGTATCTCAGTCAGTCAATGGATANGGATGACACAGAAGG 1325

2581 GAGAGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATG--ATAGGATGACAACAG-AGG 2637

1326 AAGANNAATGATCCCCACAGGCCACGACGAGTCCCACTGCTASACCCCTCAITGGGGAC 1385

2638 AAGAGAGAATGATCCCCACAGGCCACGACGAGTCCCACTGCTAGACCCCTCAITGGG-- 2694

1386 ACAGAAATCAGTAACATGGGAGATTCGTGCTCAGACACATTTGCTTAACCTGTGTCCTASAA 1445

2695 ACACAGAAATCAGAACAT- GGAGAATTGGTCTCGACACATTTGCTTAACCTGTGTCCTAGAA 2753

1446 GGACTAAGGAAAACTASGAGAAGAARTCTATYGAATTACTCAATGATGCCACCAATAACACA 1505

2754 GGACTAAGGAAAACTAGGAAG- AAGTCTTATGAATTACTCAATGATGTCCACCATAACACA 2812

1506 GGGGAAGGGAAGAAAACTCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565

2813 -GGGAAGGGAAGAAAACTCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 2871

1566 GTGCGCTCTCTGCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGCTAAGTTTATCA 1625

2872 GTGCCCTCTCTGTCCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCA 2931

1626 CTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAGTCTGCCGTAGGCCCGAGCAA 1684

2932 CTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGCTTGCCGTAGGCCCGAGCAA 2991

1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485

1685 ACTTAGAAACCCCTATTGAACCTGGCAACYTCGGTTTTATATAAGATCAGGAGGC 1744

2882

2992 ACCTAGAAACCCCTATIGAACCTGGTCTCTTATTAATAGAGATCAGGAGGAGC 3051

1745 AGGCGGAACAGGACAAACGGGATTAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG 1804

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

3052 AGCGGAACAGGACAAACGGGATTAAAAAGGCCACCGCTTAGTCATGACCCCTCAGG 3111

1805 CAAGTGGACTTTTGGAGGGCTCTGGAAAAGGGAAAAGCTGGGCCAAATTGAATGCCCTAATAGG 1864

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

3112 CAAGTGGA<sup>CTTTGGAGGCTCTGGAAAGGAAAGCTGGGCAAATTGAATGCC</sup>TAATAGG 3171

**1965 CEFTEC**

1863 GCCTGCTCCAGTGGCTACCAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAG 1924

3172 GCTTGCTTCCAGTGCGGTCTACAAGGACACTTTAAAAAGATTGTCCAAGTAGAAGTAAG 3231

1925 CCGCCCCCTTCGTC CATGCCCTTATTTCAAGGAATCACTGGAAGGCCCACTGCCCCAGG 1984

3232 CCGCCCGCTGGTCCATGCCCGCTTATTTCAAGGGAATCACTGGAAGGCCCATGCCCCAGC 3291

3232 CCCCCCCCCCCCCTATTTTCAAGGAAACACTCGCCCCCAGG 3291

1885 GGACAAAGGTCTTTTGGAGTCAGAAGCCACTAACCATGATCCAGCAGCAGGACTGAGGG 2044

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

3292 GGACAAAGGTCCTCTGAGTCAGAAAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGG 3351

2045 TGCCTGGGGCAAGCGCCATCCCATGGCCATCACCCCTCACAGAGCCCTGGGTATGCTTGACC 2104

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

3352 TGCCTGGGGCAAGCGCCATCCCATGCCATCACCCCTCACAGAGCCCTGGGTATGCTTGACC 3411

[illegible]

2105 A T G A G G G C C A G G A A G G T T G T C T C C T G G A C A C T G G T G C G G T C T T C T A G T C T T A C T C T T C 2164

3412 ATGAGGCCAGG-AGGTTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACTCTTC

.....

2165 TGTCCGGACAAC TGTCCTCCAGATCTGTCACTATTCTGAGGGGTCNTAAGACGGCA 2224

[illegible]

3471 TGGTCCGGACAACTGTCCTCCAGATCTGTCACCTA-TCTGAGGGGTCC-TAAGACGGCA 3528

2225 GTCAC TAGATACTTTTCTCCAGCCACTAAGTTATGAAC TGGGAGCTTATCTTTTCAC 2284

\_\_\_\_\_













Db 6132 TGCATCCCTGATACCTGACTCTCAATCTTGTGTGCTTGAAGATACTTCAAAACCA 6191  
Qy 2871 RCATCTCAACTCACCTGACCTRTTTTACCCCAAGGGTTGAGGATAGYCCCATCTATTT 2930  
Db 6192 ACATCTCAACTCACCTGACCTGTTTATCCCAAGGGTTGAGGATAGYCCCATCTATTT 6251  
Qy 2931 GGCAGGAGTAGCCCAAGACTGAGYCARVTMTCTATCTGACACTCTTGTCTTCRG 2990  
Db 6252 GGCAGGAGTAGCCCAAGACTGAGCAATCTCATCTCTGGACA--CTTGTCTTCGG 6309  
Qy 2991 TAGTGGAGTAGTTACTTTTTCGCTTCGCTTCAGAAACCTTGTGCAATCAAGCCACCA 3050  
Db 6310 TAGTGGAGTAGTTACTTTTTCGCTTCGCTTCAGAAACCTTGTGCAATCAAGCCACCA 6369  
Qy 3051 GCRCTCTTMAATTTCTCCGACCTGCTGCTACAGGTTTCCAAACSAARAGCTCACTC 3110  
Db 6370 GCGCTCTTCAATTTCTCGCTACCTGCTGCTACATGGTTTCCAAACCAAGGCTCAA 6429  
Qy 3111 TGCTCAGACAGGTTTAAATACCTTAGGCTTAATATTCAGAGGACCAAGGCCCTCAGT 3170  
Db 6430 TGCTCAGACAGGTTTAAATACCTTAGGCTTAAATATTCAGAGGACCAAGGCCCTCAGT 6485  
Qy 3171 GAGGAAYATCCAGCTATCTGCTTATCTCTCATCTCAATCYCAAAACCTTAAAGCACTAAGR 3230  
Db 6486 GAGGAACCAATCCAGCTTATCTGCTTATCTCTCATCTCAATCYCAAAACCTTAAAGCACTAAG 6545  
Qy 3231 GRRTTCTTGGRTAAYAGGYTTCTCCGGAATGATTTCCCGAGTTGTCGCAATAGC 3290  
Db 6546 GGATTTCTTGGGTAATAGGTTTCTCCGGAATGATTTCCCGAGTTTGGCAATAGC 6604  
Qy 3291 CAGGYCATTAATACASTAATTAAGGAACCTCAGAAAGCAATACCCATTTARTTAAGATG 3350  
Db 6605 CAGGTCATTAATACACTAATTAAGGAACCTCAGAAAGCAATACCCATTTAGTAAGATG 6664  
Qy 3351 GAYAMCTGAAGYMBRAGTGCTTCCAGGCCCTTAAAGAGGCTTAAACCCCAAGYCCCA 3410  
Db 6665 GACAACTGAAGTAGAAGTGCTTCCAGGCC-----TAAACCCAGCCCA 6710  
Qy 3411 GTGTTAAGYTTGCCAACGGGCAAGACTTTTSTTATATYRTACAGAAAAACAGRAAY 3470  
Db 6711 GTGTTAAGYTTGCCAACGGGCAAGACTTTTCTTCATATGTACAG-AAAAACAGGAAT 6769  
Qy 3471 AGCTCTRGAGTCTTTACAGRTCTCAGGAGYAGCTTTGCAACCVRTGGCRYACTGAS 3530  
Db 6770 AGCTCTAGAGTCTTTACAGATCCGAGGATGAGTTGCAACCTGTGGCGTACTGAC 6829  
Qy 3531 TAAAGAAATGATGTAGTGGAAGGTTGRCYTCATTTTAYGSGTAGTGTGGCAGT 3590  
Db 6830 TAAAGAAATGATGTAGTGGAAGGTTGACCTCATTTTACGGGTAGTGGTGGCAGT 6889  
Qy 3591 AGCAGYKTAGTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGGAC 3650  
Db 6890 AGCAGTCTAGTATCTGAAGCAGTTTAAATAATACAGGGAAGAGATCTTACTGTGGAC 6949  
Qy 3651 ATCTCATGAGTGAAAYRGATACTCATCTGCTAAAGAGACTTGTGGCTGTGACAACYCG 3710  
Db 6950 ATCTCATGATGTGAATGGCATACTCATCTGCTAAAGAGAGACTTGTGGCTGTGACAAC 7009  
Qy 3711 TTTACTTAAATCTCAGGCTTATTAAGGCGCAGTGTGCRACCTGTGCACTTGTGC 3770  
Db 7010 TTTACTTAAATCTCAGGCTTATTAAGGCGCAGTGTGCRACCTGTGCACTTGTGC 7069  
Qy 3771 AACTCTTAAACCCAGYCNCAATTTCTTCCAGACATGAAGAAAAAGATARAAYATACTGTCA 3830  
Db 7070 AACTCTTAAACCCAGCAATTTCTTCCAGACATGAAGAAAAAGATARAAYATACTGTCA 7129  
Qy 3831 ACAARTAAATTTCTAAACCTATGCCACTCGAGGGACCTTGTAGAGTTTCVTTGACTGA 3890  
Db 7130 ACAAGTAATTTCTAAACCTATGCCACTCGAGGGACCTTTTAGAGGTTTCVTTGACTGA 7189  
Qy 3891 TCCYGACCTTCAACTTGTATCTGATGGAAGTTCTTTTGTAGAAAAAGGACTTCGAAAG 3950

Db 7190 TCCGACC-TCAACTTGTATACTGATGGAAGTCTCTTTGTAGAAAAAGGACTTCGAAAAG 7248  
Qy 3951 YGGGTTATGAGTGTGATGATTAATGGAATAYTTGAAAGTAATCCCTCTCACTCCAGAA 4010  
Db 7249 TGGGTTATGCAAGTGTGATGATTAATGGAATAYTTGAAAGTAATCCCTCTCACTCCAGAA 7308  
Qy 4011 CTAGTGTCTYAGCTTGGCAGAACTAATAGCCYTCVYTKGGGCACTAGAAATAGGAGAA 4070  
Db 7309 CTAGTGTCTYAGCTTGGCAGAACTAATAGCCCTCACTTGGGCACTAGAAATAGGAGAA 7368  
Qy 4071 AAAGGAAATATATATATACAGACTCTTARTATGCTYACCTAGTCTCATGCTCCATGMRG 4130  
Db 7369 AAAGGCAATATA-ATACAGACTCTAATATGCTTACTAGTCTCTCCATGCTCCAG 7427  
Qy 4131 CAATATGSAAGAAAGGAATTTCTTACTTCYAGRGAACACCTATCMAACATCAGGAAG 4190  
Db 7428 CAATATGSAAGAAAGGAATTTCTTACTTCYAGRGAACACCTATCMAACATCAGGAAG 7487  
Qy 4191 CCATTTAGBARATTAATVATGCGTACAGAACTTARAGAGGTGCMAGCTTTACACTGCY 4250  
Db 7488 CCATTTAGBARATTAATVATGCGTACAGAACTTARAGAGGTGCMAGCTTTACACTGCC 7547  
Qy 4251 GGGGTCTATCAAAAGGAAAGGAAATASAAAGRAATYTGCCAAAGCAKATATTTGAAG 4310  
Db 7548 GGGGTCTATCAAAAGGAAAGGAAATASAAAGRAATYTGCCAAAGCAKATATTTGAAG 7607  
Qy 4311 CMAAAGAGCTGCAAGGAGGACCTTCATTTAGAAATGCTTATTAATCTTCCCTTAGTAT 4370  
Db 7608 CMAAAGAGCTGCAAGGAGGACCTTCATTTAGAAATGCTTATTAATCTTCCCTTAGTAT 7667  
Qy 4371 AGGTTATCTCCCTCGGGAACCAAGCCCTAGTACTCAGCAGGAGAAACAGATGCGGAA 4430  
Db 7668 AGGTTATCTCCCTCGGGAACCAAGCCCTAGTACTCAGCAGGAGAAACAGATGCGGAA 7727  
Qy 4431 CCTCAGAGGAGCTTTCTCCCTCGGGAACCAAGCCCTAGTACTCAGCAGGAGAAACAGAT 4490  
Db 7728 CCTCAGAGGAGCTTTCTCCCTCGGGAACCAAGCCCTAGTACTCAGCAGGAGAAACAGAT 7787  
Qy 4491 TGCTTGCAACTATCCAATGGAATACTTAAACCCCTTCATCAAAACCTTTTCACTTAGCA 4550  
Db 7788 TGCTTGCAACTATCCAATGGAATACTTAAACCCCTTCATCAAAACCTTTTCACTTAGCA 7847  
Qy 4551 TCGATAGCACCCTCARATGCGCAATCATTTATTTACTTGGACCGCCCTTTTCAAACTA 4610  
Db 7848 TCGATAGCACCCTCARATGCGCAATCATTTATTTACTTGGACCGCCCTTTTCAAACTA 7907  
Qy 4611 TCAAGCARATATKTCAGGCGCTGTGAATGTGCCARARAAATAATCCCTGCTYATCGCC 4670  
Db 7908 TCAAGCARATATKTCAGGCGCTGTGAATGTGCCARARAAATAATCCCTGCTYATCGCC 7967  
Qy 4671 AAGCTCTTTCAGGAPAAACAAAGACAGGCCATTTACCTTGRAPAACTTGGCAACTGATTT 4730  
Db 7968 AAGCTCTTTCAGGAPAAACAAAGACAGGCCATTTACCTTGGAPAACTTGGCAACTGATTT 8027  
Qy 4731 TACCCAGAGCCCAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATCTTCACT 4790  
Db 8028 TACCCAGAGCCCAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATCTTCACT 8087  
Qy 4791 GGGTTGGCAGAGGCTTTCCCTCTGAGGAGAAAGGCCCAAGAGGTAAATAAGGCACT 4850  
Db 8088 GGGTTGGCAGAGGCTTTCCCTCTGAGGAGAAAGGCCCAAGAGGTAAATAAGGCACT 8147  
Qy 4851 AGTTTCATGAAATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGGTGACAATAGCCC 4910  
Db 8148 AGTTTCATGAAATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGGTGACAATAGCCC 8207  
Qy 4911 TGCTTTCCAGGCCACAGTAAACCCAGGAGTATCCAGGCTTAGGTATACGATATCACTT 4970  
Db 8208 TGCTTTCCAGGCCACAGTAAACCCAGGAGTATCCAGGCTTAGGTATACGATATCACTT 8267  
Qy 4971 ACATCTGCGCTGAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAAYACTCA 5030  
Db 8268 ACATCTGCGCTGAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAAYACTCA 8327

QY 5031 AGGACATCTAAAAAGCAACCCAGGAAACCCACCTCAGATGCGCTGTTCTGTTGCTAT 5090  
 Db |||||  
 QY 8328 AGGACATCTAAAAAGCAACCCAGGAAACCCACCTCAGATGCGCTGTTCTGTTGCTAT 8387  
 Db |||||  
 QY 5091 AGCCTTAAAAAGAACTCTGCAACTTCCCAAAAGCAGAGCTTAGCCCATACCAAAATGCT 5150  
 Db |||||  
 QY 8388 AGCCTTAAAAAGAACTCTGCAACTTCCCAAAAGCAGAGCTTAGCCCATACCAAAATGCT 8447  
 Db |||||  
 QY 5151 GTATGGAAGCCCTTCATTAACCAATGACCTTGTGCTGACCCCAAGACAGCAACTTAGTT 5210  
 Db |||||  
 QY 8448 GTATGGAAGCCCTTCATTAACCAATGACCTTGTGCTGACCCCAAGACAGCAACTTAGTT 8507  
 Db |||||  
 QY 5211 GCAGACATCACTCTTAGCCCAATATCAACAGTTCTTAAACATTAAGGAACCTAT 5270  
 Db |||||  
 QY 8508 GCAGACATCACTCTTAGCCCAATATCAACAGTTCTTAAACATTAAGGAACCTAT 8567  
 Db |||||  
 QY 5271 CCTGTAGAAGAGGAAAGAACTATTCCACCCWGTGACATGTTATTAGTCAAGTCCCTT 5330  
 Db |||||  
 QY 8568 CCTGTAGAAGAGG--AAAGAAATATTCACCCCAAGTGACATGTTATTAGTCAAGTCCCTT 8625  
 Db |||||  
 QY 5331 CYCTCTAAATCCCATCCCTAGATATACATCTGGGAAGGACCCCTACCCAGTCAATTTATYT 5390  
 Db |||||  
 QY 8626 CCTCTAAATCCCATCCCTAGATATACATCTGGGAAGGACCCCTACCCAGTCAATTTATCT 8685  
 Db |||||  
 QY 5391 ACCCCAACTCGGTTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACTTGAGTCAAA 5450  
 Db |||||  
 QY 8686 ACCCCAACTCGGTTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACTTGAGTCAAA 8745  
 Db |||||  
 QY 5451 TCTGTGATCTGCAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTCTCTGTGAA 5510  
 Db |||||  
 QY 8746 TCTGTGATCTGCAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTCTCTGTGAA 8805  
 Db |||||  
 QY 5511 CCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAAACACGAGGAGAA 5557  
 Db |||||  
 QY 8806 CCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAAACACGAGGAGAA 8852  
 Db |||||

RESULT 8

AAS84210  
 ID AAS84210 standard; cDNA; 6394 BP.  
 XX  
 AC AAS84210;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #20014.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX

PS Claim 1; SEQ ID NO 20014; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 6394 BP; 1840 A; 1597 C; 1384 G; 1571 T; 0 U; 2 Other;

Query Match 49.9%; Score 3738.4; DB 5; Length 6394;  
 Best Local Similarity 91.1%; Pred. No. 0;  
 Matches 4026; Conservative 146; Mismatches 190; Indels 55; Gaps 21;  
 QY 1 1 CAACAATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
 Db |||||  
 QY 1959 CAAGGATTGCAATATAAACCCAGGCAATTCGAGCCAGCAACGCGCCCTTTGGGTCC 2018  
 Db |||||  
 QY 61 CTTCCCTTTGTATGGAGCTGT-----TTTCATGCTATTTCATCTCTATTAAAT 108  
 Db |||||  
 QY 2019 CTTCCCTTTGTATGGAGCTGT-----TTTCATGCTATTTCATCTCTATTAAAT 2078  
 Db |||||  
 QY 109 CTTGCAACCTG--CACTCTTCTGCTCCATGTTTCTTACGGCTCGAGTCTGCTCTCA 166  
 Db |||||  
 QY 2079 CTTGCAACCTGCTACCTCTTCTGTCAGTGTCTTTTACAGCTCTAGCTGAGCTCTCTG 2138  
 Db |||||  
 QY 167 CCGTCCACCACTGCTGTTT---GCCACCAACCGCANACCTGCGCTGACTCCCATCCCTCT 223  
 Db |||||  
 QY 2139 CAGTCCACCACTGCTGTTTGGCGCGCCACCGCAGACCGCTGCTGCTTCCATCCCTTC 2198  
 Db |||||  
 QY 224 GGATCCTGAGGGTGTCCGCTGTGCTCTGATCCAGCGARGCGCCCATTCGCGCTCCCAA 283  
 Db |||||  
 QY 2199 GGATCCGGCAGGGTGTCTGCTGCAATTCCTGATCCAGCGAGGCGCCCATTCGCCACGCCAA 2258  
 Db |||||  
 QY 284 TTGGGCTAAAGGCTTGCCATTTGCTGTCAGCGCTAAGTCCCTGGGTTGTTCTAATGA 343  
 Db |||||  
 QY 2259 TCGGGCTAAAGGCTTGCCATTTGCTGCAAGTTAAGTGGCGGGTTGGTCTGTAATGA 2318  
 Db |||||  
 QY 344 GCTGAACACTANTCACTGGGTTCCATGTTCTCTCTGTGACCCACGCGCTCTTAATA-KA 402  
 Db |||||  
 QY 2319 GCTGAACACTGCTCACTGGGTTCCAGGTTCTCTTCCGTGACCCCATGCTCTAATAGA 2378  
 Db |||||  
 QY 403 ACTATAACACTTACACAT-GGCCCAAGATTCCATTCTT-GGAATCCGTGAGGCAACG 460  
 Db |||||  
 QY 2379 ACTATTAAACATTACACATGGGCCCAAGATTCCATTCTCTTGGGAATTCGCTGAGGCCAAG 2438  
 Db |||||  
 QY 461 AACTCCAGGTGAGAAATACGARGCTTGCCACCATCTTGGAGCGGCTGCTACCTCTT 520  
 Db |||||  
 QY 2439 AACCCCGGTTCAGAGATATCAAGGGTGTGCCACCATCTTGGAGCGGCGCTGCCATCTT 2498  
 Db |||||  
 QY 521 GGAAGTGTTCACCACTTCTGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTG 580  
 Db |||||  
 QY 2499 GGAACGGCTTGCACCATCTTGGAGCTCTGTGCACAGGACCCCGGTGAACTTTG 2558  
 Db |||||  
 QY 581 GCRACCAACGAGCATCCMAAGTGTGGGAAAGCTTCCCGCAAGACAAAGCGCCC 640  
 Db |||||

Db 2559 GCAACCCAGAACGGACATCAAAGTGATGGAAAGCGTTTCCCGCAAGACAAAAACGCCCC 2618  
Qy 641 TAAGACGTATTTCTGABAAATTTGGGAMCAATTTGACCTTCAGACACTAA-GAAAGAAACGA 699  
Db 2619 TAAGACGTATTTCTGGAGAAATTTGGGACCAATTTGACCTTCAGACACTAAAGGAAAGAAACGA 2678  
Qy 700 CTTATATTTCTTCTGCAGTGGCCCTCGGCACCTCTGAGGAGAGATATAAATTAATTAACACCAT 759  
Db 2679 CTTATATTTCTTCTGCAGTGGCCCTCGGCACCTCTGAGGAGAGATATAAATTAATTAACACCAT 2738  
Qy 760 CTTACAGCTAGACACTTTCTTTGTAGAA---AAGGCAAAATGGAGTGCATTAAGTACA 816  
Db 2739 CTTACAGCTAGACACTTTCTTTGTAGAAAGAGAGTAAATGGAGTGAAGTCTATATGTACA 2798  
Qy 817 AACTTTCTTTTCAATTAAGAGACAACTCAAAATTA-TGTAAAAAGTGTGATTTATGCC-- 873  
Db 2799 AACTTTCTTTTCAATTAAGAGACAACTCAAAATTAATTTGTAAGAGTGTGATTTATGCCCT 2858  
Qy 874 TACAGGAAGCCTTCAGAGTCTACCTCCCTATCCC---AGCATCCCGACCTCTTCCCA 929  
Db 2859 ACAGGGAAGCCTTCAGAGTCTACCTCCCTATCCCGACATCCCCCAACTCTTCTCCA 2918  
Qy 930 MYTAATAAGGACCCCTTCAACCCAAATGGTCCAAAAGAGAGATAGACAAAAGGGTAAAC 989  
Db 2919 ACTATTAAGGACCCCTTCAACCCAAATGGTCCAAAAGAGAGATAGACAAAAGGGTAAAC 2978  
Qy 990 AGTGAACAAAGAGTGCCTAATTTCCCAATTAATGACCCCTCCCAAGCAGTGGGAGGAAG 1049  
Db 2979 AGTGAACAAAGAGTGCCTAATTTCCCAATTAATGACCCCTCCCAAGCAGTGGGAGGAAG 3038  
Qy 1050 AGAATTCGGCCAGCAGAGTGCATGTCYTTTTTCTCCAGACTTAAAGCAATAAA 1109  
Db 3039 AGAATTCGGCCAGCAGAGTGCATGTCYTTTTTCTCCAGACTTAAAGCAATAAA 3098  
Qy 1110 AACAGACTTAGTAAATTTCTCAGATAAATTCGATGGCTATATGRTGTTTTTACAGGGTT 1169  
Db 3099 AACAGACTTAGTAAATTTCTCAGATAAATTCGATGGCTATATGRTGTTTTTACAGGGTT 3158  
Qy 1170 AGGACAAATTTCTGATCTGACATGGAGAGATATATGTCATGCTGCTAAATCAGACACTAA 1229  
Db 3159 AGGACAAATTTCTGATCTGACATGGAGAGATATATGTCATGCTGCTAAATCAGACACTAA 3218  
Qy 1230 CCCCAATGAGAGAGTCCCACTAACTGACGCTGAGRTTGGCGATCTCTGGTATC 1289  
Db 3219 CCCCAATGAGAGAGTCCCACTAACTGACGCTGAGRTTGGCGATCTCTGGTATC 3278  
Qy 1290 TCAGTCAGTCAATGGATANGATGACAAACAGAGAAAGANAATGATTTCCCAAGGCC 1349  
Db 3279 TCAGTCAGTCAATGGATA-GGATGACAAACAG-AGANAAGAGACGATTTCCCAAGGCC 3336  
Qy 1350 AGCARGCAGTTCCTAGCTTASACCTCTATTTGGGACACAGAAATCAGTAACATGGAGAT 1409  
Db 3337 AGCARGCAGTTCCTAGCTTASACCTCTATTTGGGACACAGAAATCAGTAACATGGAGAT 3396  
Qy 1410 TGGTCTGCAGACATTTGCTAATCTGTGTCTASAGGACTAAGGAAACTASGAGAAA 1469  
Db 3397 TGGTCTGCAGACATTTGCTAATCTGTGTCTAAGGACTAAGGAAACTACGAGAAA 3456  
Qy 1470 RCTAYGAATTAATCAATGATGTCCACCAATAACAGGGGAAGGAAAGAAATCTCTACTG 1529  
Db 3457 ATCTAGCAATTAATCAATGATGTCCACCAATAACAGGGGAAGGAGANAATCTCTACTG 3516  
Qy 1530 CTTTCTGGAGAGACTAAGGGAGGCAATTTAGGAAGCGTCTCTCTGTCACCTGACTCTT 1589  
Db 3517 CTTTCTGGAGAGACTAAGGGAGGCAATTTAGGAAGCGTCTCTCTGTCACCTGACTCTT 3576  
Qy 1590 CTGAGGCCCACTAATCTTAAGCGTAAAGTTTATCACTCAGTCAGTCAGCAGCATTTAGNA 1649  
Db 3577 CTGAGGCCCACTAATCTTAAGCGTAAAGTTTATCACTCAGTCAGTCAGCAGCATTTAGNA 3636  
Qy 1650 AAAAATTCAAAAGTCTGCGCTAGGCCCGGAGCAAAAATTTAGAAACCTTATTTGAACTTGGC 1709  
Db 3637 AAAAATTCAAAAGTCTGCGCTAGGCCCGGAGCAAAAATTTAGAAACCTTATTTGAACTTGGC 3696

Qy 1710 AACYTCGGTTTTTTATATAGAGATCAGGAGGACAGCGGAAACAGGACAAACCGGATTA 1769  
Db 3697 AACYTCGGTTTTTTATATAGAGATCAGGAGGACAGCGGAAACAGGACAAACCGGATTA 3756  
Qy 1770 AAAAAAGGCGCACCGCTTTTAGTCAATGACCCCTCAGGCAAGTGGACTTTTGAGGGCTCTCGAA 1829  
Db 3757 AAAAAAGGCGCACCGCTTTTAGTCAATGACCCCTCAGGCAAGTGGACTTTTGAGGGCTCTCGAA 3816  
Qy 1830 AAGGGAAGCTGGGCAAAATTTGAATGCTTAATAGGGCTTGTCTCAGTGCGGTCTTACAAG 1889  
Db 3817 AAGGGAAGCTGGGCAAAATTTGAATGCTTAATAGGGCTTGTCTCAGTGCGGTCTTACAAG 3876  
Qy 1890 GACACTTTAAAAAGATTTGTCCAAGTAGAAGTAAAGCGCCCTTCGTCATGCCCTTAT 1949  
Db 3877 GACACTTTAAAAAGATTTGTCCAAGTAGAAGTAAAGCGCCCTTCGTCATGCCCTTAT 3936  
Qy 1950 TTCAAGGGAATCACTGGAAGGCCCACTGCCCAAGGGAACAAAGGTCTTTTGAAGTCAGAAG 2009  
Db 3937 TTCAAGGGAATCACTGGAAGGCCCACTGCCCAAGGGAACAAAGGTCTTTTGAAGTCAGAAG 3996  
Qy 2010 CCACCTAACAGATGATCCAGCAGCAGGACTGAGGGTGTCTGGGGCAAGCGCATCCCATG 2069  
Db 3997 CCACCTAACAGATGATCCAGCAGCAGGACTGAGGGTGTCTGGGGCAAGCGCATCCCATG 4056  
Qy 2070 CCATCACCTCACAGAGCCCTGGGTATGCTTGACCATTTAGGGCCAGGAAGGTGTCTCC 2129  
Db 4057 CCATCACCTCACAGAGCCCTGGGTATGCTTGACCATTTAGGGCCAGGAAGGTGTCTCC 4116  
Qy 2130 TGGACACTGGTGGCTCTTCTTAGTCTTACTTCTTCTGTCGGGACAACTGTCTCTCCAGAT 2189  
Db 4117 TGGACACTGGTGGCTCTTCTTAGTCTTACTTCTTCTGTCGGGACAACTGTCTCTCCAGAT 4176  
Qy 2190 CTGTCACTATTCTCAGGGGGTCCNTAAGACGGGCAGTCACTAGATPACTTTTCTCCAGCA 2249  
Db 4177 CTGTCACTATTCTCAGGGGGTCC-TAAGACGGGCAGTCACTAGATPACTTTTCTCCAGCA 4235  
Qy 2250 CTAAGTTATGAACTGGGGAGCTTTTATCTTTTACATGCTTTTCTTAATATGCTTGAAG 2309  
Db 4236 CTAAGTTATGAACTGGGGAGCTTTTATCTTTTACATGCTTTTCTTAATATGCTTGAAG 4295  
Qy 2310 CCCCACTACCTTGTAGGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTTATACACTGAA 2369  
Db 4296 CCCCACTACCTTGTAGGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTTATACACTGAA 4355  
Qy 2370 CATAGGAAGGAACACCCGTTTGTGTNCCCTGCTTGTAGGAAGGAATTAATCTGAAAG 2429  
Db 4356 CATAGGAAGGAACACCCGTTTGTGT-CCCTGCTTGTAGGAAGGAATTAATCTGAAAG 4414  
Qy 2430 TCTGGGCAACAGAGGACAAATATGAGCAGGCAAAAGATGCGCTCTGTTCAAGTTAAA 2489  
Db 4415 TCTGGGCAACAGAGGACAAATATGAGCAGGCAAAAGATGCGCTCTGTTCAAGTTAAA 4474  
Qy 2490 CTAAGGATTCACCTTCTCTTCCCTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAA 2549  
Db 4475 CTAAGGATTCGCTCTCTTCCCTACCAAGGACATACCCCTTAGACCTGAGGCCCAA 4534  
Qy 2550 CAA--GGATTCCAAAAGATTGTTAAGGACTTTAAAGCCCAAGGGCTTAGTAAAAACCATGCA 2607  
Db 4535 CAAAGGGCTTCCAAAAGATTGTTAAGGACTTTAAAGCCCAAGGGCTTAGTAAAAACCATGCA 4594  
Qy 2608 TAACTCCCTGCAAGTAATTCCTAGTGGATTTGAGGAGGACAGAAACCCAGTGGACAGTGG 2667  
Db 4595 TAACTCCCTGCAAGTAATTCCTAGTGGATTTGAGGAGGACAGAAACCCAGTGGACAGTGG 4654  
Qy 2668 AGGGTTAGTCAAGATCTCAGGATTTATCAATGGAGGCCGTGTCTTTTATATACCCAGCTG 2727  
Db 4655 AGGGTTAGTCAAGATCTCAGGATTTATCAATGGAGGCCGTGTCTTTTATATACCCAGCTG 4714  
Qy 2728 TACCTAGCCCTTATCTGTGMYTTCCCAATATCCAGAGGAAGCAGAGTGGTTTTTACSTCC 2787  
Db 4715 TACCTAGCCCTTATCTGTGCTTTCCCAATATCCAGAGGAAGCAGAGTGGTTTTTACACTCC 4774







XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 4349 BP; 1319 A; 1061 C; 941 G; 1026 T; 0 U; 2 Other;

Query Match 45.8%; Score 3431; DB 5; Length 4349;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 3974; Conservative 148; Mismatches 195; Indels 237; Gaps 28;

QY	115	ACTGCACCTCTTCTGCTCCATGTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCAC	174
DB	1	ACTGCACCTCTTCTGCTCCATGTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCAC	60
QY	175	CACCTGCTGTTTGGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCTCTGCAG	234
DB	61	CACCTGCTGTTTGGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCTCTGCAG	120
QY	235	GGTGTCCGCTGTGCTCTGTATCCAGGA---RGGCCCATTTGCCGCTCCCAATGGGCTA	291
DB	121	GGTGTCCGCTGTGCTCTGTATCCAGGA---RGGCCCATTTGCCGCTCCCAATGGGCTA	180
QY	292	AAGGCTTGCCATTTGTCCTGAC--GGCTAAGTGCT--GGTGTGTTCTTAATTGAGCTGAA	349
DB	181	GAGGCTCACCATTGTTCTGACAGGCTAAGTGCTGCGGTTTCACTTAATCGAGCTGAA	240
QY	350	CACCTANTCACTGGGTTCCATGGTTCTCTTCTGTGACCCCAACCGCTTCTAATAKAACTATA	409
DB	241	CACCTAATCACTGGGTTCCATGGTTCTCTTCTGTGACCCCAACCGCTTCTAATAKAACTATA	300
QY	410	CACCTACCAATGGCCCAAGATTCCATCTCTTGGAAATCCGTGAGGSCAACGNACTCCAGG	469
DB	301	CACCTACCAATGGCCCAAGATTCCATCTCTTGGAAATCCGTGAGGSCAACGNACTCCAGG	359
QY	470	TCAGAGAATACGAGCTTGCCACCATTTGGAAGCGGCTGTACCTCTTGGAAAGTGT	529
DB	360	TCAGAGAATACGAGCTTGCCACCATTTGGAAGCGGCTGTACCTCTTGGAAAGTGT	419
QY	530	TCACCAACCATCTTGGAGCTCTGTGAGCAAGGACCCCGGCTTACATTTTGGSCRACCA	589
DB	420	TCACCAACCATCTTGGAGCTCTGTGAGCAAGGACCCCGGCTTACATTTTGGSCRACCA	479
QY	590	RACGACATCCMAAGTATGGGAAAGTTTCCCGCAAGACAAACACCGCCCTTAAGACGTA	649
DB	480	GACGACATCCMAAGTATGGGAAAGTTTCCCGCAAGACAAACACCGCCCTTAAGACGTA	539
QY	650	TTCTGGARAATTGGGAAAGTTTCCCGCTCAGACACTAAGAAGAAACGACTTATATCT	709
DB	540	TTCTGGARAATTGGGAAAGTTTCCCGCTCAGACACTAAGAAGAAACGACTTATATCT	599
QY	710	TTCTGAGTCCGCTGGCACTCTCTGAGGGAAGTATATAATTAACCACTTTACAGCTA	769
DB	600	TTCTGAGTCCGCTGGCACTCTCTGAGGGAAGTATATAATTAACCACTTTACAGCTA	659

QY	770	GACYTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTCCATAGTACAACTTTCTTTTCA	829
DB	660	GACCTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTCCATAGTACAACTTTCTTTTCA	719
QY	830	TTAAGAGCAACTCACAAATATATGTAAGAGTGTGATTTATCCCTTACAGGAAGCTTCA	889
DB	720	TTAAGAGCAACTCACAAATATATGTAAGAGTGTGATTTATCCCTTACAGGAAGCTTCA	779
QY	890	AGTCTACTCTCCCTATCCAGCATCCCGACTCCCTTCCCACTTAAATAGGACCCCTTC	949
DB	780	AGTCTACTCTCCCTATCCAGCATCCCGACTCCCTTCCCACTTAAATAGGACCCCTTC	839
QY	950	AACCCAAATGTCCTAAGAGGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCAA	1009
DB	840	AACCCAAATGTCCTAAGAGGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCAA	899
QY	1010	TATTTCCCAATATGACCCCTTCCCAAGCAGTGGGAGGAGAGAATTCGGCCCGACGAG	1069
DB	900	TATTTCCCAATATGACCCCTTCCCAAGCAGTGGGAGGAGAGAATTCGGCCCGACGAG	958
QY	1070	TGCATGTGCTTTTCTTCTCCAGACTTAAAGCAATTAACCAAGTGTAGTAAATTTCT	1129
DB	959	TGCATGTGCTTTTCTTCTCCAGACTTAAAGCAATTAACCAAGTGTAGTAAATTTCT	1018
QY	1130	CAGATAATCTGATGGCTATATGTTTTCAGGGTTAGGACAAATTTCTTGTATCTGA	1189
DB	1019	CAGATAATCTGATGGCTATATGTTTTCAGGGTTAGGACAAATTTCTTGTATCTGA	1078
QY	1190	CATGGAGAGATATATGTCCTCTAAATCAGACACTAACCCCAAAATGAGAGAGTGCC	1249
DB	1079	CATGGAGAGATATATGTCCTCTAAATCAGACACTAACCCCAAAATGAGAGAGTGCC	1137
QY	1250	ACCATTAATGAGCTGAGGTGTTGGCGATCTCTGTATCTCAGTCAGTCAATGGATAN	1309
DB	1138	ACCATTAATGAGCTGAGGTGTTGGCGATCTCTGTATCTCAGTCAGTCAATGGATAN	1196
QY	1310	GGATGACAAAGAGGAAAGANAATGATTTCCCAAGCAGCAGCAGCAGTTCCTCAGTCTA	1369
DB	1197	GGATGACAAAGAGGAAAGANAATGATTTCCCAAGCAGCAGCAGCAGTTCCTCAGTCTA	1256
QY	1370	SACCTCATTTGGGGACACAGAAATCAGTAAATGGGAGATTTGGTCTGCAGACATTTGCT	1429
DB	1257	SACCTCATTTGGG---ACACAGAAATCAGAAAT---GGAGATTTGGTCTGCAGACATTTGCT	1312
QY	1430	AACCTGTGTCTASAGGACTAAGGAAACTTASGAGAAATCTTAYGAATTTACTCAATGA	1489
DB	1313	AACCTGTGTCTASAGGACTAAGGAAACTTASGAGAAATCTTAYGAATTTACTCAATGA	1371
QY	1490	TGTCACCATTAACAGGGGAGGAGAAATCTTCTGCTTCTTCTGGAGAGACTAAGG	1549
DB	1372	TGTCACCATTAACAGGGGAGGAGAAATCTTCTGCTTCTTCTGGAGAGACTAAGG	1430
QY	1550	GAGGCATTTGAGGAGCGTCTCTCTGTCTCAGTCTCTCTCTCTGAGGCGCACTAATCTTA	1609
DB	1431	GAGGCATTTGAGGAGCGTCTCTCTGTCTCAGTCTCTCTCTCTGAGGCGCACTAATCTTA	1490
QY	1610	AAGCGTAAGTTTATCATCTCAGTCTCAGATTCAGATTAAGAAACTTCAAAAGTCTGCG	1669
DB	1491	AAGCGTAAGTTTATCATCTCAGTCTCAGATTCAGATTAAGAAACTTCAAAAGTCTGCG	1550
QY	1670	TAGCCCGGAGCAAACTTAGAAACCTTATTTGAACTTCGGCACTTCGGTTTTTATTAATA	1729
DB	1551	TAGCCCGGAGCAAACTTAGAAACCTTATTTGAACTTCGGCACTTCGGTTTTTATTAATA	1610
QY	1730	GAGATCAGGAGGAGCAGCGGAAACAGGACAAACCGGATTTAAATAAGGCGCAACCGCTTA	1789
DB	1611	GAGATCAGGAGGAGCAGCGGAAACAGGACAAACCGGATTTAAATAAGGCGCAACCGCTTA	1670
QY	1790	GTCTGATCCCTCAGGCAAGTGTCTTTGGAGGCTCTTGGAAAGGAGGAAAGCTGGGCAAT	1849
DB	1671	GTCTGATCCCTCAGGCAAGTGTCTTTGGAGGCTCTTGGAAAGGAGGAAAGCTGGGCAAT	1730
QY	1850	TGAATGCTTAATAGGCTTGTCTTCCAGTGTGCTTCTAAGGAGACTTTTAAATAAGATTTGT	1909

Db 1731 TGAATGCTTAATAGGCTTCTTCCAGTGGGTCTACAGGACACTTTAAAAAGATTGT 1790  
Qy 1910 CCAAGTAGAAGTAAAGCCGCTTCTGTCATGCCCTTTATTTCAAGGAATCACTGGAAG 1969  
Db 1791 CCAAGTAGAA----- 1800  
Qy 1970 GCCCACTGCCCCAGGGGACAAAGTCTTTTGGAGTCAAGAGCCACTAACCCAGATGATCCAG 2029  
Db 1801 ----- 1800  
Qy 2030 CAGCAGACTGAGGGTGCCTTGGGGCAAGCGCCATCCCATGCCATCCCTCACAGAGCCC 2089  
Db 1801 -----AGCCC 1805  
Qy 2090 TGGGTATGCTTGACCAATTGAGGGCCAGGAAGGT-----TGTCTCTGGACACTGGTGGGTC 2146  
Db 1806 CAGGTATGCTTGACCAATTGAGAGCCAAAGAGTTAACTGTCTCTGGACGCGGTGGCGCC 1865  
Qy 2147 TTCTTAGTCTTACTCTTCTGTCGCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGG 2206  
Db 1866 TTCTCAGTCTTACTCTCTGTCGCGGACAACTGTCTCCAGATCTGTCACTATCC---GA 1922  
Qy 2207 GGGTCNTAAGAGCGGCGAGTCACTAGATACTTTTCCAGCCACTAAGTTTATGAATGGG 2266  
Db 1923 GGGGCTTAGGACAGCCAGTCACTAGATAC-TTCTTCCAGCCACTAAGTTGTG-ACTGGG 1980  
Qy 2267 GAGCTTTATTTTTCACATGCTTTTCTAATATGTTTGAAGCCCACTACCTTGTAG 2326  
Db 1981 GAGCTTTTACCTTTTTCACATGCTTTCTAATATGCTTGAAGCCCTCACTCCCTTGTAG 2040  
Qy 2327 GGAGAGACATTTAGCAAGAGCGGGCCATTATACCTGAACATAGGAGAGGAACAC 2386  
Db 2041 AGAGAGAGATTTAGTAAAGAGCGGGCCATTATACCTGAACATAGGAGAGGAACAC 2100  
Qy 2387 CCGTTTGTGTGTCCTGCTTGAAGAGGAATTAATCTCGAAGTCTGGGCAACAGAGGA 2446  
Db 2101 CCAATTTGTGCT-CTCCTGTTTGAAGAGGAAATTAGTCTTGAAGTCTGGGCAACAGAGGA 2159  
Qy 2447 CAATATGGAGCGCCAAAGAAATCCCGTCTGTTTCAAGTTTAACTAAAGGATTTCCACTTC 2506  
Db 2160 CAATATGGCAAG-CAAGAATGCCATTTTGTCCAAGTTTAACTAAAGGATTTCCACCTC 2218  
Qy 2507 CTTTCCCTACCAAGAGGAGTACCCCTCAGACCCAAAGGCCCAACAGAGATTCCAAAGAT 2566  
Db 2219 CTTTCCCTATCAAGAGCAGTACCCCTCAGACCCGAGGCCCAACAAAGATTCCAAAGAT 2278  
Qy 2567 TGTAAAGGACTTAAAGCCCAAGGCTTAGTAAACCATGCATACCTCCCTGCGAGTAATTC 2626  
Db 2279 TGTTAAGGACCTAAAGCCCAAGGCTTAGTAAACCATGCATACCTCCCTGCGAGTAATTC 2338  
Qy 2627 CGTATGGATTTGAGGAGGACAGAAAACCCAGTGGACAGTGGAGGGTTAGTGCAGATCTC 2686  
Db 2339 -----AATTTTAGGATACAGAAACCCAAAGACAGTGA-GGTTAGTGCAGATCTC 2390  
Qy 2687 AGGATTTAATAGGAGGCGGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTATACTGT 2746  
Db 2391 AGGATTTAATCAAT-GAGGCTGTTTCTCTATACCCAGCTATACCTAGCCCTTATACTCT 2449  
Qy 2747 GWTTCCTCAATACAGAGAGCAGAGTGGTTTACSTCTTGACCTTMAAGATGCTT 2806  
Db 2450 GCTTTTCCCAATACAGAGAGCAGAGTGGTTTACAGTCTCGGATCTTAAAGGATGCTT 2509  
Qy 2807 CTTCTGCATCCCTGTATCATCTGACTCTCAATTTCTTTGTTGCTTTTGAAGATACTTCAAA 2866  
Db 2510 TTTCTGCATCCCTGTATCATCTGACTCTCAATTTCTTTGTTGCAATTTGAAGATCTTCAAA 2569  
Qy 2867 CCCATCATCTCAACTCACTGGACTTTTTTACCCCAAGGGTTTCAAGGATAGYCCCATCT 2926  
Db 2570 CCCAATCTCAACTCACTGGACTGTTTTTACCCCAAGGGTTTCAAGGATAGTCCCATCT 2629  
Qy 2927 ATTTGGCCAGGATTAAGCCCAAGACTTGAGYCAATWTATACCTGGACACTCTTGTCT 2986

Db 2630 ATTTGGCCAGGCAATTAGCCCAAGACTTTGAGCCAGTTCTCATACCTGGACACTCTCTGTCT 2689  
Qy 2987 TCRGTATGCTGGATGATTTATCTTTTTRGCGCYRVTTCAGAAACCTTGTGCATCAAGCCAC 3046  
Db 2690 TCAGTGCATGGATGATTTATCTTTTAGTCCCGTTTCAAGAACCTTGTGCGCATCAAGCCAC 2749  
Qy 3047 CCAAGCRCTCTTMAATTTCTCGCYACCTGTGCTACAMGGTTTCCAAACSBARAGTCA 3106  
Db 2750 CCAAGCGCTCTTAAATTTCTCACTACCTGTGCTCAAGGTTTCCAAACCAAGGCTCA 2809  
Qy 3107 RCTCTGCTCACAGCAGTTTAAATFACITTAGGRCTTAARATTTATCCAAAGGACCACARGGCCCT 3166  
Db 2810 GCTCTGCTCACAGCAGTTTAAATGCTTAGGCTTAAATTTATCCAAAGTCAACAGGCGCT 2869  
Qy 3167 CAGTGAAGAAATYATCCAGCCTATACCTGCTTATCTCATCTCYCAAAACCTTAAAGCAACT 3226  
Db 2870 CAGTGAAGAAACGTTATCCAGCCTATACCTGCTTATCTTATCCAAACCTTAAAGCAACT 2929  
Qy 3227 AAGRGRRTTCTTGGCCTAAAYAGGTTTCTGCCGAAATATGGAAT-CCAGGTACGGCAAA 3286  
Db 2930 AAGAGGTTTCTTGGCATAACAGGTTTCTGCCGAAATATGGAAT-CCAGGTACGGCAAA 2988  
Qy 3287 TAGCCAGGYCATTAATACATAAATTAAGGAAACTCAGAAAGCCAAATACCCATTTARTAA 3346  
Db 2989 TAGCCAGACGGTTATATACGCTTAATTAAGGAAACTCAGAAAGCCAAATACCCATTTAGTAA 3048  
Qy 3347 GATGAYAMCTGAAGYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCAAGY 3406  
Db 3049 GATGGACACTGAAGCAAGCAGCTTTCCATGCCCTTAAAGAGGCGCT--AATGCAAGC 3106  
Qy 3407 CCGAGTTTAAGYTTGGCCAAAGGGGCAAGACTTTTCTCATATGTCACAAAAAACAACA 3454  
Db 3107 CCGAGTGTAAAGTGTCTTAAACGGGGCAAGACTTTTCTCATATGTCACAAAAAACAACA 3166  
Qy 3455 -----AGAAAAAACAAGAAVAGCTCTGGAGTCTTACACAGRTCCRAGGAYGA 3505  
Db 3167 AACAAAGAAAAAACAACATGATAGCTCTAAGAGTCTCTGGCTGGTCTGAGGACGA 3226  
Qy 3506 GCTTGCACCCYRTGGCRVACCTGASTAAGGAAATGATGTAGTGCAAAAGGGTTGRCYTC 3565  
Db 3227 GTTTGCAACCCCTGGCATACCTGA-TAAGGAAATGATGTAGTGCAAAAGGGTCAGCCTC 3285  
Qy 3566 ATTTGTTTAGGTAGTGGTGGCAGTAGCAGTGTAGTATCTGAAGCAGTTAAATTAATAC 3625  
Db 3286 ATTTGTTTAGGTAGTGGTGGCAGTAGCAGTGTAGTATCTGAAGCAGTTAAATTAATAC 3345  
Qy 3626 AGGAGAGAGATCTTACTGTGCGACATCTCATGAGTGGAAYRGCATACCTCACTGCTAAG 3685  
Db 3346 AGGAGAGAAATCTTACTGTGCGACATCTCATGATGTGAATGGCATACTCACTGCTAAG 3405  
Qy 3686 GAGACTTGTGGCTGTGAGCAACACCTTTTACTTAAATRTCAGGCTCTATTTACTTGAARGGC 3745  
Db 3406 GAGACTTGTGGCTGTGAGCAACACCTTTTACTTAAATATCAGGCTCTATTTACTTGAAGGC 3465  
Qy 3746 CAGTGTGCRATGTGCACTTGTGCAACTCTTTAAACCCAGYCNATTTCTTCCAGACAAATG 3805  
Db 3466 CAATGTGCGACTGTGCACTTGTGCACTTCTTAAACCTGAGACACATTTCTTCCACAAATG 3525  
Qy 3806 AAGAAAGATARAAYATAACTGTCAACAAATTTTCTCAAAACCTATGCACTCGAGGGG 3865  
Db 3526 AAGAAAGATARAAYATAACTGTCAACAAATTTTCTCAAAACCTATGCACTCGAGGGG 3585  
Qy 3866 ACTTTTAGAGTGTCTTGTGACTGTATCCAGCTTCCAGCTTCAACTGTATATCTCATGAGTCTC 3925  
Db 3586 ACCATTTAGAGGTTCCCTTGACTGTATCCACCCCT-ACTTGTATATCTGATGGAAGATCC 3644  
Qy 3926 TTTGTAGAAAAAGGACTTCGAAAAAGYGGGTTATGCAAGTGGTCACTGATTAATGGAATYTT 3985  
Db 3645 TTTGTAGAAAAAGGACTTCGAAAAAGTGGGTTATGCAAGTAGTCACTGATTAATGGAATYTT 3704  
Qy 3986 GAAAGTAAATCCCTCACTCCAGGAACTAGTGTCTGAGTGGCAAGAACTAATAGCCYTCAT 4045  
Db 3705 GAAAGTAAATCCCTCACTCCAGGAACTAGTGTCTGAGTGGCAAGAACTAATAGCCYTCAT 3755



Db 4463 TCACCACCATCTTTGGGAGCTCTGTGAGCAAGGACCCCGGTAAACATTTTGGCAACCAAC 4522  
Qy 590 RACGACATCCMAAGTAGTGGGAAAAGCTTCCCGCAAGACAAAAGCCCGCTTAAGACGTA 649  
Db 4523 GACGGACATCCAAAGTAGTGGGAAAAGCTTCCCGCAAGACAAAAGCCCGCTTAAGACGTA 4582  
Qy 650 TTCTGGARAAATTTGGGAMCAATTTGACCCCTCAGACACTTAAGAAAAGAACGACTTATATCT 709  
Db 4583 TTCTGGAGAAATTTGGGACCAATTTGACCCCTCAGACACTTAAGAAAAGAACGACTTATATCT 4642  
Qy 710 TCTGCAAGTCCGCTGGCACTCTCTGAGGGAAAGTATAAATTTAACCACCATCTTTACAGCTA 769  
Db 4643 TCTGCAAGTCCGCTGGCACTCTCTGAGGGAAAGTATAAATTTAACCACCATCTTTACAGCTA 4702  
Qy 770 GACVCTCTTTGTAGAAAAGCAATGGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCA 829  
Db 4703 GACCTCTTTGTAGAAAAGCAATGGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCA 4762  
Qy 830 TTAAGAGACAACTCACAAATTTATGTAAGAAAGTGTGATTTATGCCCTACAGGAAGCCCTTCAG 889  
Db 4763 TTAAGAGACAACTCACAAATTTATGTAAGAAAGTGTGATTTATGCCCTACAGGAAGCCCTTCAG 4822  
Qy 890 AGTCTACCTCCCTATCCAGCATCCCGACTCTCTTCCCACTTAAAGGACCCCGCTTC 949  
Db 4823 AGTCTACCTCCCTATCCAGCATCCCGACTCTCTTCCCACTTAAAGGACCCCGCTTC 4882  
Qy 950 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAA 1009  
Db 4883 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAA 4942  
Qy 1010 TATTTCCCAATATGACCTCCCAAGCAGTGGGAGGAGAGAAATTTGGGCCAGCCAGAG 1069  
Db 4943 TATTTCCCAATATGACCTCCCAAGCAGTGGGAGGAGAGAAATTTGGGCCAGCCAGAG 5002  
Qy 1070 TGCATGTGCVTTTTYTTCTCCAGACTTAAGCAAAATAAAGCAAGTCTTGTGTAATCTT 1129  
Db 5003 TGCATGTGCVTTTTCTCTCCAGACTTAAGCAAAATAAAGCAAGTCTTGTGTAATCTT 5062  
Qy 1130 CAGATAAATCTGATGCTATATGRTGTTTAAAGGGTTAGGACAAATCTTTGTACTGA 1189  
Db 5063 CAGATAAATCTGATGCTATATGRTGTTTAAAGGGTTAGGACAAATCTTTGTACTGA 5122  
Qy 1190 CATGGAGATATATGTCACGTCTAATCAGACACTAACCCCAATCAGAGAAAGTGCC 1249  
Db 5123 CATGGAGATATATGTCACGTCTAATCAGACACTAACCCCAATCAGAGAAAGTGCC 5182  
Qy 1250 ACCATAAATCTGAGCTGAGGTTTGGCGATCTCTGCTATCTCAGTCAGGTCATGATGATAN 1309  
Db 5183 ACCATAAATCTGAGCTGAGGTTTGGCGATCTCTGCTATCTCAGTCAGGTCATGATGATAN 5241  
Qy 1310 GGATGACAAAGAGAAAGAAATGATTTCCCAAGGCGAGCAGTCAGTCCAGTCTA 1369  
Db 5242 GGATGACAAAGAGAAAGAG-ATGATTTCCCAAGGCGAGCAGTCAGTCCAGTCTA 5300  
Qy 1370 SACCCTCATTTGGGACACAGAAATCAGTAAATGGAGATTTGCTGTCAGACATTTGCT 1429  
Db 5301 GACCCCTCATTTGGGACACAGAAATCAGTAAATGGAGATTTGCTGTCAGACATTTGCT 5360  
Qy 1430 AACTTGTGTCTASAGGACTAAGGAAACTASAGAAATCTAAGAAATCTAATGATCAATGA 1489  
Db 5361 AACTTGTGTCTAAGGAACTAAGGAAACTAAGAAATCTAAGAAATCTAATGATCAATGA 5420  
Qy 1490 TGTCCACCAATAACAGGGGAAGGAAAGAAATCTTCTGCTTCTGAGAGACTTAAGG 1549  
Db 5421 TGTCCACCAATAACAGGGGAAGGAAAGAAATCTTCTGCTTCTGAGAGACTTAAGG 5480  
Qy 1550 GAGGCATTTGAGGAAGCGCTCTCTGTCACCTGACTCTTCTGAGGCGCAACTTAATCTTA 1609  
Db 5481 GAGGCATTTGAGGAAGCGCTCTCTGTCACCTGACTCTTCTGAGGCGCAACTTAATCTTA 5540  
Qy 1610 AAGCGTAAGTTTATCTAGTCAGTCAGACATTTAGAAAAAACTTCAAAAAGTCTGCCG 1669  
Db 5541 AAGCGTAAGTTTATCTAGTCAGTCAGACATTTAGAAAAAACTTCAAAAAGTCTGCCG 5600

Qy 1670 TAGGCCCGGAGCAAAAATTAGAAAACCTATTGAACTTGGCAACCTCGTTTTTTTATAATA 1729  
Db 5601 TAGGCCCGGAGCAAAAATTAGAAAACCTATTGAACTTGGCAACCTCGTTTTTTTATAATA 5660  
Qy 1730 GAGATCAGGAGGAGCAGGCGGACAGACAAAAGGATTAAGAAAAGGCGCCGCTTTA 1789  
Db 5661 GAGATCAGGAGGAGCAGGCGGACAGACAAAAGGATTAAGAAAAGGCGCCGCTTTA 5720  
Qy 1790 GTCATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTGAAAAGGAAAAGCTCGGCAAAAT 1849  
Db 5721 GTCATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTGAAAAGGAAAAGCTCGGCAAAAT 5780  
Qy 1850 TGAATCCCTAATAGGGCTTCTTTCAGTGGCGTCTAAGAGACACTTTTAAAAAGATTTGT 1909  
Db 5781 TGAATCCCTAATAGGGCTTCTTTCAGTGGCGTCTAAGAGACACTTTTAAAAAGATTTGT 5940  
Qy 1910 CCAAGTAGAAGTAAGCCGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAG 1969  
Db 5841 CCAAGTAGAAGTAAGCCGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAG 5900  
Qy 1970 GCCCACTGCCCGGAGCAAAAAGTCTTTTGAAGTCAAGAGCCCACTAACAGATGATCCAG 2029  
Db 5901 GCCCACTGCCCGGAGCAAAAAGTCTTTTGAAGTCAAGAGCCCACTAACAGATGATCCAG 5960  
Qy 2030 CAGCAGGACTGAGGGTGCCTGGGCAAGCGCCATCCCATGCCATCACTCCCTCACAGAGCCC 2089  
Db 5961 CAGCAGGACTGAGGGTGCCTGGGCAAGCGCCATCCCATGCCATCACTCCCTCACAGAGCCC 6020  
Qy 2090 TGGGTATGCTTGACCACTTGAGGGCCAGGAAAGTGTCTCTGAGACACTGGTGGGCTTTC 2149  
Db 6021 TGGGTATGCTTGACCACTTGAGGGCCAGGAAAGTGTCTCTGAGACACTGGTGGGCTTTC 6080  
Qy 2150 TTAGTCTTACTCTCTGTCGCCGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGG 2209  
Db 6081 TTAGTCTTACTCTCTGTCGCCGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGG 6140  
Qy 2210 TCNTAAGACGGGCGAGTCACTAGATACTTTTCCCAGCCCACTAAAGTTATGAACTGGGGAG 2269  
Db 6141 TCC-TAAGACGGGCGAGTCACTAGATACTTTTCCCAGCCCACTAAAGTTATGAACTGGGGAG 6199  
Qy 2270 CTTTATCTTTTTCATGCTTTTCTAATTTATGCTTGAAGAGCCCACTACTCTGTTTAGGGA 2329  
Db 6200 CTTTATCTTTTTCATGCTTTTCTAATTTATGCTTGAAGAGCCCACTACTCTGTTTAGGGA 6259  
Qy 2330 GAGACATTTAGCAAAAGCAGGGCCATTAACCTGAACATAGAGAGGAACACCCG 2389  
Db 6260 GAGACATTTAGCAAAAGCAGGGCCATTAACCTGAACATAGAGAGGAACACCCG 6319  
Qy 2390 TTTGTTGTNCCCTCTGTTGAGGAAGGAATTAATCTGAAAGTCTGGGCAACAGAGGACAA 2449  
Db 6320 TTTGTTGT-CCCTCTGTTGAGGAAGGAATTAATCTGAAAGTCTGGGCAACAGAGGACAA 6378  
Qy 2450 TATGACGAGCAAAAGATGCCCCGCTCTGTTCAAGTTAACTAAAGGATTTCACTTCCCT 2509  
Db 6379 TATGACGAGCAAAAGATGCCCCGCTCTGTTCAAGTTAACTAAAGGATTTCACTTCCCT 6438  
Qy 2510 TCCCTACCAAGGCGAGTACCCCTCAGACCCAGGCGCCCAAGGATTTCCAAAGATTTGT 2569  
Db 6439 TCCCTACCAAGGCGAGTACCCCTCAGACCCAGGCGCCCAAGGATTTCCAAAGATTTGT 6498  
Qy 2570 TAAAGCACTTAAAGGCCAAGGCTTAGTAAACCATGATTAACCTCTGAGTAATTTCCCT 2629  
Db 6499 TAAAGCACTTAAAGGCCAAGGCTTAGTAAACCATGATTAACCTCTGAGTAATTTCCCT 6558  
Qy 2630 AGTGAATTTGAGGAGCAGACAGAAACCCAGTGGACAGTGGAGGGTGTAGTGAAGATCTCAG 2689  
Db 6559 AGTGAATTTGAGGAGCAGACAGAAACCCAGTGGACAGTGGAGGGTGTAGTGAAGATCTCAG 6618  
Qy 2690 ATTATCAATTTGAGGCGGTTGCTCTTTTATACCCAGTGTACCTAGCCCTTATCTGTGMY 2749  
Db 6619 ATTATCAATTTGAGGCGGTTGCTCTTTTATACCCAGTGTACCTAGCCCTTATCTGTGCT 6678











CC in the specification. A polynucleotide of the invention has cytostatic  
CC activity, and may have a use in gene therapy, or in a vaccine. The  
CC recombinant nucleic acids and polypeptides are useful for treating  
CC carcinomas, e.g. lymphomas, cancers, neoplasms, adenocarcinoma, and  
CC sarcomas. The present sequence represents a human gene of the invention.  
XX  
SQ Sequence 22436 BP; 7692 A; 4872 C; 4296 G; 5576 T; 0 U; 0 Other;

Query Match	40.7%;	Score 3049.8;	DB 10;	Length 22436;
Best Local Similarity	86.5%;	Pred. No. 0;		
Matches 3556;	Conservative 142;	Mismatches 320;	Indels 93;	Gaps 31;
QY	606	GATGGGAAACGTTCCCGCAGACAAAAAAGCGCCCTAAGAGTATTTCTGGABAATTCGGGA	665	
DB	1581	GGTAGCAAAACATTCCTCTAAGGCAGAAACACCCCGAAGATGTATTTCTGGAGAATCGGGA	1640	
QY	666	MCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATATTTCTTCGCAAGTGCAGCTG	725	
DB	1641	CCAATTTGACCTTCAGACGCTGAGAAAGAAATGACTTATATTTCTTCGCAATTTGCCTG	1700	
QY	726	GC-----	753	
DB	1701	GCACAAATATCTCTTCATGGGGAGAAACGTGGCCCTCTCGAGGAAAGTATAAATATATGA	1760	
QY	754	CACCATCTTACAGCTAGACACTTTTTCGTAGAA--	810	
DB	1761	CACCATCTTACAGCTAGACCTCTTTTGTAGAAAGAGGCNAATGGAGTGAAGCGCCATA	1820	
QY	811	AGTACAAACTTTCTTTTCATTAAGAGACAACTCACAAATATGTAAAAAGTGTGATTTATG	870	
DB	1821	TGTACAAATTTTCTTTTCATTAAGAGACAACTCGCTATTATGTAAAAAGTGTGATTTATC	1880	
QY	871	CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCAGCAT--	928	
DB	1881	CCCTACAGGAAGCCCTCAGAGTGACCTTCTTACCCAGCATCCCCCTACTCTTCCCC	1940	
QY	929	AMTAAATAAGGACCCCCCTTCAACCCAAATGTGTCAAAAGGAGATAGACAAAAGGGTAAA	988	
DB	1941	AACATAAGGACCCCCCTTCAACCCAAATAGTCCAAAGAGAGATAGACAAAAGGGTAAA	2000	
QY	989	CAGTGAACCAAAGAGTGCCTAATATTCCTCCAAATATGACCCCTCCAGCAGTGGGAGGA	1048	
DB	2001	CAATGAATCAAGAGTGCCCAATATTTCTCCGATTATGCCCCACT--	2059	
QY	1049	GAGAAATTCGGCCAGCCAGAGTGCATGTGCTTTTCTCCAGACTTAAAGCAATAA	1108	
DB	2060	GAGAAATTCGGCCAGCCAGAGTGATGATCCATTTCTCTCTCAGTCTTAAAGCAATATG	2119	
QY	1109	AAACAGACTTAGGTAAATTCCTCAGATAATCTCATATGCTATATTTGRTGTTTTTCAAGGTT	1168	
DB	2120	AAATAGACCTAGGTAAATTTCTAGCTTAACCTCATGGCTATATTTGATTTTAAAGGGT	2179	
QY	1169	TAGGACAAATCTTTTGATCTGACATGGAGAGATATATGTCACTGCTTAATCAGACACTA	1228	
DB	2180	TAGGACAAATCTTTTGATCTGACATGGAGAGATATA-ATGTTACTGTTAAATCAGACACTA	2238	
QY	1229	ACCCCAATAGAGAGAGTGCACCAATAATCTGACGCTGAGRGTTTGGCGATCTCTGGTAT	1288	
DB	2239	ACCCCAATAGAGAGAGTGCAGCGCAATAACTGACGCGGAGAGTTTGGCGATCTCTGGTAT	2298	
QY	1289	CTCAGTCAGGTCAATGGATANGGATGACACAGAGGAAGANAATGATTTCCCCACAGGC	1348	
DB	2299	CTCAGTCAGGTCCAT--CATAGGTGAAACAG-AGGAAGAGACGATTTCCCAAGGC	2355	
QY	1349	CAGCARGAGTTCCTCAGTCTAASACCTCATTTGGGGACACAGAAATCAGTAAACATGGGAGA	1408	
DB	2356	CAGCARGAGTTCCTCAGTGTAGACCTCTCATAGGG--	2411	
QY	1409	TTGCTGCTGCAGACATTTGCTAACTTGCTGTASAGGACTAAGGAAACTAGAGAA	1468	
DB	2412	ATGGTCAGCAGACATTTGCTAACTTGCTGTGTAGAGGACTAAGGAAACTAGGAAG-A	2470	
QY	1469	ARTCTAYGAATTACTCAATGATCTCAACCATAAACACAGGGGAAGGAAAGAAATCTCTACT	1528	

Db	2471	AGCCTATGAATTAATTCATGATGCTCCAGTATAACA-AAGGGAAGAAGAAAATCCTACT	2529
Qy	1529	GCCTTTCTGGAGAGACTAAGCGAGGCATTCAGGAAGCGTCCCTCTCTGTCACTGACTCT	1588
Db	2530	GCCTTTCTGGAAGACTAAGGGAGGCATTCAGGAGGAAACCTGTCTGTCACTGACTCT	2589
Qy	1589	TCTGAAGGCCAACTAATCTTTAAAGCGTAAGTTATATCATCTCAGTCTAGCTGCAGACATTAG-	1647
Db	2590	TTTGAAGGCCAACTAATCTTTAAAGGAAAGTTTATCACTCAGTCTCAGCTGCAGACATTAGA	2649
Qy	1648	AAAAAAGCTTCAAAGTCTCGCGTAGGCGCGGAGCAAACTTAGAAACCTTATGAACTTG	1707
Db	2650	AAAAAAGCTTCAAAGTCTCGCGTAGGCGCGGAGCAAACTTAGAAACCTTATGAACTTG	2709
Qy	1708	GCAACTYTCG--TTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGGAACAGGACAAACGGGA	1766
Db	2710	GCAACTCGGTTTTTTATAATAGAGTCAAGGAGGAGCAGGCGGGAATGGGACAAACGGGA	2769
Qy	1767	TTAAAAAAGGCCACCGCTTTTAGTATGATGACCTCAGGCAAGTGGACTTTGGAGGCTCTG	1826
Db	2770	-TAAAAAAGGCCACCGCTTTTAGTATGATGACCTCAGGCAAGTGGACTTTGGAGGCTCTG	2828
Qy	1827	GAAAGGGAAAAGCTGGGCAATTTGAATGCTTAATAGGCGTCTTCCAGTGGCGTCTAC	1886
Db	2829	GAAAGGGAAAAGCTGGGCAATTTGAATGCTTAATAGGCGTCTTCCAGTGGCGTCTAC	2888
Qy	1887	AAGGACACTTTAAAAAGATTTGTTCAAGTAGAAGTAAAGCGGCCCTTCGTCCATGCCCCCT	1946
Db	2889	AAGGACACTTTAAAAAGATTTTCCAAGTACAATTAAGCCACCCCTCGTCCATGCCCCCT	2948
Qy	1947	TATTTCAAGGGAATCTCTGGAAGGCCACATGCCCCAGGGGACAAAGGTCTTTTGAGTCTAG	2006
Db	2949	TATATCAAGGGAATCTCTGGAAGGCCACAGGGCCCCGGGGATGAAGGTCTCTGTAGTCTAG	3008
Qy	2007	AAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGGCCATCCC	2066
Db	3009	AAGCCACTAACAGATGATCCAAACAGCAGGACTGAGGGTTCCTGGGGCAAGCACCAGCCC	3068
Qy	2067	ATGCCATCACCCCTCACAGAGCCCTGGGTATGCTTTGACCATTGAGGGCCAGGAAGGT--T	2123
Db	3069	ATGCCATCACCCCTCATAGAGCCCTGGGTATGCTTTGACCTTTGAGGGCCAGGAGGTTAACT	3128
Qy	2124	GTCTCTGGACACTGGTGGCGTCTCTTAGCTTTACTTCTGTCTCCGGACAACTGTCTCT	2183
Db	3129	GTCTCTGGACACTGGCAGGCCCTTCTCAGTCTTACTGTCTCTGTCTGTGACAACTGTCTCT	3188
Qy	2184	CCAGATCTGTCACTATTCTGAGGGGGTCTNTAAGACGGGAGTCACTAGATCTTTTCTC	2243
Db	3189	CCAGATCTGTCACTATCC--GAGGGGTCTTAGGACAGCAGTCACTAGATATCTTCTC	3244
Qy	2244	CAGCCACTAAGTTATGAACTGGGAGCTTTTATCTTTTCCATGCTTTTCTAATTAATGCT	2303
Db	3245	CAGCCACTAAGTTATG-CTGGGGAATTTTACTCATTTTCCATGCTTTTCTAATTAATGCT	3303
Qy	2304	TGAAGCCCCACTACCTTTGTAGGAGAGACATTTCTAGCAAAGCAGGGGCCATTATACA	2363
Db	3304	TGAAGCCCCACTCTCTTTGTAGTGAAGAAATTTCTAGCAAAGCAGGGGCCATTATACA	3363
Qy	2364	CCTGAACATAGGAAAGGAACACCCGTTTGTGTGTCCTGTTTGGGAAGGAATTAATC	2423
Db	3364	CCTGAACATAGGAAAGGAACACCCGTTTGTGTGT-CCCCCTGTTGAGGAAGGAATTAATC	3422
Qy	2424	CTCAAGTCTGGGCACAGAAAGCAATATGGAGCGCCAAAGATGCCCGTCTCTGTCAA	2483
Db	3423	CTCAAGTCTGGGCACAGAAAGCAATATGGAGCGCTATGGATGAGC--AAAGAAATGCCAT	3480
Qy	2484	GTTAAACTAAGGATTTCCACTTCTTTTCCCTACCAAAGGCAGTACCCCTCAGACCCAG	2543
Db	3481	GTTAAACTAAGGATTTCTGCTCTTCTTCTTACCAAAGGCAGTA-CCCCTTAGACCCAG	3539
Qy	2544	GCCCAACAAAGGATTTCAAAGATTTGTTAAGGACTTTAAAGCCCAAGGCTTAGTAAACCA	2603

Db 3540 GCCCAACAAGGACTCCAAAAGATTGTTAAGGACCGCAAAAGCCCAAGGCTCTAGTAAACCA 3599  
Qy 2604 TGCATACTCCCTGCAAGTAATTCCTGAGTGGATTGAGGAGGCACAGAAACCCAGTGGACA 2663  
Db 3600 TGCATAAGCCCTTGCAATATTC-----AATTTTAGAGTACAGAAACCCAGTGGACA 3652  
Qy 2664 GTGGAGGGTTAGTGCAGAGATCTCAGAGATTATCAATGGAGGCGTGTCTCTTTATACCCA 2723  
Db 3653 GTGGA-GGTGTAGTGCAGAGATCTCAGAGATTATCGAT-TAGGCTGTGTGTCTCTATATCA 3710  
Qy 2724 GCTGTACCTAGCCCTTATATCTGTGMYTTTCCCAAATACCAGAGAAAGCAGAGTGGTTTACA 2783  
Db 3711 GCTGTACCTAACCCCTTATACCTGCTTTCCCAAATACCAGAGAAAGCAGAGTGGTTTACA 3770  
Qy 2784 STCTGGACCTTMAAGATGCCCTTCTCTGCAATCCCTGTACATCTCTGACTCTCAATCTTGTG 2843  
Db 3771 GTCTCTGGACCTTAAGGATGTCTTTTCTGTCATCCCTGTACATCTCTGACTCTCAATTTTTA 3830  
Qy 2844 TTTGCTTTTGAAGATPACTTCAAAACCCARCATCTCAACTCACCTGAGACTTTTACCCCAA 2903  
Db 3831 TTTGCTTTTGAAGATCTTTTGAACCAATGTCTCAACTCACCTGAGACTGTTTTACCCGAG 3890  
Qy 2904 GGGTTCCAGGATAGYCCCCATCTATTTGGCCAGGCAATTAGCCCAAGACTTGAGYCAATYM 2963  
Db 3891 GGGTTCCAGATAGCCCCCATCTATTTGGCCAGGCAATTAGCCCAAGACTTGAGCCAAATTC 3950  
Qy 2964 TCATACCTGGACACTTGTCTCTTCRGTAKGTGGATGATTTACTTTTTRGCGCCVRRTTCA 3023  
Db 3951 TCATACCTGGACACTTGTCTCTTGGTACATGSGATGATTTACTTTTAGCCACCCATTTCA 4010  
Qy 3024 GAAACCTTGTGCATCAAGCCACCAAGCRCTTTMAATTTCTCGVACCTGTGGCTAC 3083  
Db 4011 GAAGCTTTGTGCATCAAGCCACCAAGCACTTTAAATTTCTCGCTACCTGTGGCTAC 4070  
Qy 3084 AMWGTTTCCAAASARAGCTCARCTCTGCTCACAGCAGGTTTAAATACTTAGGRTCAARA 3143  
Db 4071 AA-GATTCAAACCAAGGCTCAGCTCTGCTCACAGAGTTTAAATACTTAGGCGCTAAA 4129  
Qy 3144 TTATCAAAGGCCAARGGCCCTCAGTGAAGAAAYATCCAGCTTATCTTGCTTATCTCT 3203  
Db 4130 TTATCAAAGGTACCAAGGCCCTCAGTGGGGAACATATGCAAGCTTATCTTGCTTGTCT 4189  
Qy 3204 CATYCNAAACCTTAAGCAACTAAGRBRTTCTTGGCTTAAYAGGYTTCTGCCGAWA 3263  
Db 4190 CATCCAAAACCTT-AAGCAACTAAGAGGGTTCTTGGCATAAACAGGTTTCTGCCAATA 4248  
Qy 3264 TGGATTCCCAGGTWTGGCRAAATAGCCAGGYCATTTAWATACASTAAATTAAGGAACTCA 3323  
Db 4249 TGGATT-CCAGGTATGGCAAAATAGCCAGATCATTTATACACTTAATTAAGGAACTCA 4307  
Qy 3324 GAAAGCCAAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAGTGGCTTTCCAGGCCCC 3383  
Db 4308 GAAAGCCAAATACCCATTTAGTAAAGTGGACACCTGAAGCACAAGCAGCTTTCCAGGCTCT 4367  
Qy 3384 TAAAGAGCCCTTAACCCCAAGYCCAGTGTAAAGYTTTGGCAACGCGGCAAGACTTTTST 3443  
Db 4368 AAAGAA--GGCTTAACCCCAAGTTCAGTGTACGCTTGCCAAAGGGGCAAGATTTTCT 4425  
Qy 3444 TVATAYRTCACAGAAAAAACAAGRAAYAGCTCTRGAGTCTTTACACAGRTCCRAGGGAY 3503  
Db 4426 TTAATGTTCACAGAAAAAACAAGAAATAGCTCTAGAGTCTTTACACAGGTCCGAGGAT 4485  
Qy 3504 GAGCTTGCACCYRTGGCRYACTGASTAAGAAAYTGATGTAGTGGCAAGGGTGTGRY 3563  
Db 4486 GAACTTGGCAACTGTGGCATACTGAGTAAGGAAATTTGATGCAGTGGCAAGGGTGTGCC 4545  
Qy 3564 TCATTTGTTTAYGGGTAGTGGCGAGTAGCAGTYKTATGTATCTGAAGCAGTTAAATATAT 3623  
Db 4546 TCATTTGTTTATGGGTAGTGGCGAGGTAGCAGTCTTAGTATCTGAAGCAGTTAAATATAT 4605  
Qy 3624 ACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCAATCTCACTGCTAA 3683  
Db 4606 ACATGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGCGCATACTCACTGCTAA 4665

Qy 3684 AGGAGACTTGTGGCTGTCCAGACAAACYGTTTACTTTAAATRTCCAGGCTCTATTACTTTGAARG 3743  
Db 4666 AGGAGACTTGTGGCTGTCCAGACGATCGTTTACTTTAAACATCATGAGGCTCTATTACTTTGAARG 4725  
Qy 3744 GCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTTAAACCAGYCNCAATTTCTTCCAGACAA 3803  
Db 4726 GCCTGTGTGCAACTGTGCGGCTTTGCAACTCTTTAAACCAGGCAATTTCTTCCAGACAA 4785  
Qy 3804 TGAAGAAAGATARAAYATAAATCTGTCAACAARTAAATTTCTCAAACTATGCCACTTCGAGG 3863  
Db 4786 TGAAGAAAGATAGAACAATACTGTCAACAAGTAATTTGCTCAAACTACCCGCTCGAGG 4845  
Qy 3864 GGACCTTGTAGGCTTCTTGTGATCCYGCATCCYGCATCTCAACTTGTGTATCTGATGGAAGTT 3923  
Db 4846 GGACCTTTTAGAGGTTCCCTTGTGATCTCCGACTTCCGACTT-AACTGTGTATCTGATGGAAGTT 4904  
Qy 3924 CTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGCAGTGCTCAGTGTATGATTAATGGAATAY 3983  
Db 4905 CTTTGTAGAAAAAGGACTTTGAAA-----AGTGTGAGGATTAATGGAATATC 4952  
Qy 3984 TTGAAGTAAATCCCTCTCACTCCAGAACTAGTGTCTAGTGTGCAAGAACTAATAGCCYTC 4043  
Db 4953 TTGAATAATAATCCCTCTCACTCCAGAACTAGTGTCTCAGCCGCAATACTAATAGCCCTCA 5012  
Qy 4044 YTKGGCACTAGATTAAGGAGAGRAAAAGGYYAATATATATACAGACTCTTRATATG 4103  
Db 5013 CTCGAGCACTAGAAATTTAGGAGAGAAAGGYYAATATATATACATAGACTCTTAAGTATG 5072  
Qy 4104 CTYACCTAGTCNTCCATGCCCATMRGCAATATATGARAAGAAAGGAAATCTTAACTTCVG 4163  
Db 5073 CTTACCTAGTCTCTCATGCCCCATGAGCAATATGGAGAGAAAGGAAATCTTAACTTCTG 5132  
Qy 4164 AGRGAACACTTATCAMACATCAGGAAGCCATTAAGARATTAATTAATGTCGWTGACAGAAC 4223  
Db 5133 AGGGAACACTTATCAACCATCAGGAAGCAATTAAGAGATTAATTAATGCTGTACAGAAC 5192  
Qy 4224 CTARAGAGTGTGMACTTCTTACCTGCTGCGGCTCATCANAAAGAAAGAAAGGAAATAS 4283  
Db 5193 CCNAAGAGGTGGCAGTCTTACACTGCTAGGCTCATCAGAGAAGAAAGGAAAGGAAATAG 5252  
Qy 4284 AAGRNAATGCAAGCAKATATTAAGCMAAAGAGCTGCAAGGAGGAGCCCTCCATTAAG 4343  
Db 5253 AAGGAACCAACAGCAGATTAATGAGCCAAAGAGCTGCAAGACAGGACCCCTCCATCAG 5312  
Qy 4344 AAATGCTTTATTAACCTTCCCTTAGTATAGGTAATCCCTTCCGGGAAACCAAGCCCCAGT 4403  
Db 5313 AAATGCTTTATAGAGGACCCCTTAGTGTGGGTAAACCCCTCCAGGAAACCAAGCCCCAGT 5372  
Qy 4404 ACTCAGCAGGAGAAACAGAAATGGGAACTCACAGG--CAGTTTCTCCCTCGGGAC 4460  
Db 5373 ACTCAGCAGGATGAATAGAAATGGGAACTCATGAGGACATATCTTCTCCCTGAGGAT 5432  
Qy 4461 GGTAGCCACTGAAGAGGAAATATCTTTGCTGCAACTATCAATGGAATTAATTTA 4520  
Db 5433 GCCTAGCCCAAGAGAGAAATATCTTTGCTGCACTAACCAATGGAATTAATTTA 5492  
Qy 4521 AAACCTTTTATCAAACTTTCACTTAGGCACTGATAGCACCCTCATATGCGCAATCAT 4580  
Db 5493 AAACCTTTTATCAAACTTTTCACTTTAGACATTTGATAGCACCCTCATATGCGCAATTTAT 5552  
Qy 4581 TATTACTGGACAGGCTTTTCAAACTATCAAGCAATATCAGGCGCTGTGAATGT 4640  
Db 5553 TATTACTGGACAGGCTTTTCAAACTATCAAGCAATATCAGGCGCTGTGAATGT 5612  
Qy 4641 GCCARAAAAATATCCCTGCTTATCGCCA 4671  
Db 5613 GCCAAGAAATAACCCCTGCTGCTGAGGCCA 5643

RESULT 13  
ADC85361  
ID ADC85361 standard; DNA; 22436 BP.

XX ADC85361;  
 AC 01-JAN-2004 (first entry)  
 DT Mouse Pap coding sequence.  
 XX  
 DE Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;  
 KW secreted; transmembrane; intracellular; ds.  
 XX  
 XX Mus sp.  
 OS  
 XX WO2003045230-A2.  
 PN  
 XX 05-JUN-2003.  
 PD  
 XX  
 XX 02-DEC-2002; 2002WO-US038582.  
 PF  
 XX 30-NOV-2001; 2001US-00997722.  
 PR  
 XX (SAGR-) SAGRES DISCOVERY.  
 PA  
 XX Morris DW, Engelhard EK;  
 PI WPI; 2003-513603/48.  
 DR  
 XX  
 XX New recombinant nucleic acid comprising a nucleotide sequence of any of  
 PT the carcinoma-associated (CA) genes, useful for screening for drug  
 PT candidates for diagnosing or treating carcinomas.  
 PT  
 XX Claim 1; SEQ ID NO 147; 983bp; English.  
 PS  
 XX The invention relates to a recombinant nucleic acid comprising a  
 CC nucleotide sequence selected from any of the fully defined carcinoma-  
 CC associated (CA) genes from the 50 tables given in the specification. The  
 CC CA proteins are secreted, transmembrane or intracellular proteins. The  
 CC recombinant nucleic acids are useful for screening for drug candidates  
 CC for diagnosing or treating carcinomas. Sequences given in ADC85215-  
 CC ADC8514 represent CA genes of the invention.  
 CC  
 XX SQ Sequence 22436 BP; 7692 A; 4872 C; 4296 G; 5576 T; 0 U; 0 Other;  
 Query Match 40.7%; Score 3049.8; DB 10; Length 22436;  
 Best Local Similarity 86.5%; Pred. No. 0;  
 Matches 3556; Conservative 142; Mismatches 320; Indels 93; Gaps 31;  
 QY 606 GATGGAAACGTTCCCGCAAGACAAAAACGCCCTTAAGACGTTATCTCGAATAATGGGA 665  
 DB 1581 GGTAGCAACATTCCTCCCTCAAGCAAAAAACCCCGAAGATGTTATCTCGAATAATGGGA 1640  
 QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAAGCACTTATATCTCTGCGAGTCCGCGCTG 725  
 DB 1641 CCAATTTGACCTCAGAGCTGAGAAAGAAATGACTTATATCTCTGCGAGTATGCGCTG 1700  
 QY 726 GC-----ACTCTGAGGGAAGTATAAATATAA 753  
 DB 1701 GCCACAATATCTCTTATGGGGGAGAAACGTTGGGCTCTGAGGGAAGTATAAATATAA 1760  
 QY 754 CACCATCTTACAGTAGACTCTTTTGTAGAA---AAGCAATGGAGTGAAGTCCATA 810  
 DB 1761 CACCATCTTACAGTAGACTCTTTTGTAGAAAGAAAGCAATGGAGTGAAGCGCCATA 1820  
 QY 811 AGTACAAACTTTCTTTTCAATTAAGAGACAACTCACAATATATGTAAGAAAGTGTATTATG 870  
 DB 1821 TGTACAAATTTCTTTTCAATTAAGAGACAACTCGCTATTTATGTAAGAAAGTGTATTATC 1880  
 QY 871 CCTTACAGAAAGCTTACAGTCTACCTCCCTATCCAGCAT--CCCGACTCTTTCCCG 928  
 DB 1881 CCTTACAGAAAGCTTACAGTCTACCTCCCTATCCAGCAT--CCCGACTCTTTCCCG 1940  
 QY 929 AMYTAATAGGACCCCTTCAACCAATGTCCAAAGGAGATAGACAAAAGGGTAAA 988  
 DB 1941 AACTAATAGGACCCCTTCAACCAATGTCCAAAGGAGATAGACAAAAGGGTAAA 2000

QY 989 CAGTGAACCAACAGAGTCCCAATATTTCCCAATATGACCCCTCCCAAGCAGTGGAGGAA 1048  
 DB 2001 CAATGAATCAACAGAGTCCCAATATTTCCCAATATGACCCCTCCCAAGCAGTGGAGGAG 2059  
 QY 1049 GAGAAATTCGGCCCGCAGCAGAGTGCATGTCGTTTTTCTCCAGACTTTAAAGCAATAA 1108  
 DB 2060 GAGAAATTCGGCCCGCAGCAGAGTGCATGTCACCAATTTCTCTCTCAGTCTTAAAGCAATA 2119  
 QY 1109 AAACAGACTTAGGTAAATTTCTCAGATAAATCTCAGATAAATCTGATGGCTATATTTGTTTCAAGGGT 1168  
 DB 2120 AAATAGACCTTAGGTAAATTTCTCAGCTAAACCTGATGGCTATATTTGTTTCAAGGGT 2179  
 QY 1169 TAGGACAATTTCTTGTATCTGACATGAGAGATATATGTCTACTGCTTAATCAGACACTA 1228  
 DB 2180 TAGGACAATTTCTTGTATCTGACATGAGAGATATA-ATGTTACTGTTAAATCAGACACTA 2238  
 QY 1229 ACCCAATGAGAAAGTGCACCACTAATCTGACCTGAGAGTGTGGCGATCTCTGGTAT 1288  
 DB 2239 ACCCAATGAGAAAGTGCACCACTAATCTGACCTGAGAGTGTGGCGATCTCTGGTAT 2298  
 QY 1289 CTCAGTCAGTCAATGGATANGATGACAAACAGAAAGAAAGAAATGATTTCCCAACAGGC 1348  
 DB 2299 CTCAGTCAGTCAAT--CATAGATGAAACACAG-AGAAAGAGAACGATTTCCCAACAGGC 2355  
 QY 1349 CAGCAGCAGTTCCTCAGTCTASACCTCATTTGGGACACAGAAATCAGTAAACATGGGAGA 1408  
 DB 2356 CAGCAGCAGTTCCTCAGTCTAGACCTCATAGGG---ACACAGATCAGACAT-AGAGA 2411  
 QY 1409 TTGGTGTCTCAGACATTTCTAACTTGTGTCTASAGGACTAAGGAAACTASGAGAA 1468  
 DB 2412 ATGTGTCTCAGACATTTCTAACTTGTGTCTAGAGGACTAAGGAAACTAGGAAG-A 2470  
 QY 1469 ARTCTAYGAATTTACTCAATGATGTCCACCATACACAGGGGAGGAGAAATCTCTACT 1528  
 DB 2471 ACCTATGAATTTTCAATGATGTCCAGTATAACA-AAGGAAAGGAGAAATCTCTACT 2529  
 QY 1529 GCCTTTCTGAGAGACTAAGGAGGAGTGTAGAGGAGCGTGCCTCTCTGTCACTGACTCT 1588  
 DB 2530 GCCTTTCTGAGAGACTAAGGAGGAGTGTAGAGGAGGAAACCTGTCTCACTGACTCT 2589  
 QY 1589 TCTGAGGCAACTAATCTTAAAGCGTAAAGTTATACCTCAGTCAGTCTCAGACATTAG- 1647  
 DB 2590 TTTGAGGCAACTAATCTTAAAGGAAAGTTTATCACTCAGTCAGTCTCAGACATTAGA 2649  
 QY 1648 AAABAACTTCAAAAGTCTGCGGTAGCGCCGAGGAGGAAACTTGAACCCCTTATGAATTG 1707  
 DB 2650 AAAAAGCTTCAAAAGTCTGCGGTAGCGCCGAGGAGGAAACTTGAACCCCTTATGAATTG 2709  
 QY 1708 GCAACVTGG-TTTTTTATATAGAGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1766  
 DB 2710 GCAACVTGG-TTTTTTATATAGAGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2769  
 QY 1767 TTAATAAAAGGCCACCGCTTTAGTCAATGACCTCAGGCAAGTGGACTTTGGAGGCTCTG 1826  
 DB 2770 -TAAATAAAAGGCCACCGCTTTAGTCAATGACCTCAGGCAAGTGGACTTTGGAGGCTCTG 2828  
 QY 1827 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1886  
 DB 2829 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2888  
 QY 1887 AAGGACACTTTTAAATAAAGATTTGTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1946  
 DB 2889 AAGGACACTTTTAAATAAAGATTTGTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2948  
 QY 1947 TATTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2006  
 DB 2949 TATTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3008  
 QY 2007 AAGCAGCTAATCAGATGATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2066  
 DB 3009 AAGCAGCTAATCAGATGATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3068



Db 5193 CCAAAGAGTGGCAGTCTTACACTGCTAGGTCATCAGAGAAAGGAAGGAATAG 5252  
 QY 4284 AAGGAAATGCCAAGCAKATATTGAACMAAAGAGCTGCAAGCAGGACCCCTCCATTAG 4343  
 Db 5253 AAGGAAATGCCAAGCAGATATTGAAGCCAAAGAGCTGCAAGCAGGACCCCTCCATTAG 5312  
 QY 4344 AAATGCTTATTAACCTTCCCTAGTATAGGTAATCCCTTCCGGAACCAAGCCCAAGT 4403  
 Db 5313 AAATGCTTATAGAGGACCCCTAGTGTGGGGTAACCCCTTCCAGGAACCAAGCCCAAGT 5372  
 QY 4404 ACTCAGCAGGAGAAACAGAAATGGGAACCTTCAAGAGG--CAGTTTTCTCCCTCGGGAC 4460  
 Db 5373 ACTCAGCAGGATGAATAGATGGGAACCTTATGAGGACATATCTTCTCCCTGAGGAT 5432  
 QY 4461 GGTAGGCACTGAAGAGGAAATATCTTTGCTGCAATATCCATGCAATGAATATCTTA 4520  
 Db 5433 GGCTAGCCACCAAGAGGAAATATCTTTGCTGCAATATCCATGCAATGAATATCTTA 5492  
 QY 4521 AAACCTTTCATCAACCTTTCACCTAGGATGATAGGACCCATCAGATGGCCCAATCAT 4580  
 Db 5493 AAACCTTTCACCAAACTTTCACCTAGGATGATAGGACCCATCAGATGGCCCAATCAT 5552  
 QY 4581 TATTACTGACCAAGGCTTTTCAAACTATCAAGCAATATCAAGGCTGTGAATGT 4640  
 Db 5553 TATTACTGACCAAGGCTTTTCAAACTATCAAGCAATATCAAGGCTGTGAATGT 5612  
 QY 4641 GCCAARAATAATCCCTGCTTATCGCA 4671  
 Db 5613 GCCAARAATAATCCCTGCTTATCGCA 5643

RESULT 14

ADMT74477  
 ID ADM74477 standard; DNA; 22436 BP.  
 AC ADM74477;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human carcinoma associated (CA) nucleic acid #73.  
 XX  
 KW Human: carcinoma associated nucleic acid; CA nucleic acid; gene; ds;  
 KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;  
 KW cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004072154-A1.  
 XX  
 PD 15-APR-2004.  
 XX  
 PF 30-NOV-2001; 2001US-00997722.  
 XX  
 PR 22-DEC-2000; 2000US-00747377.  
 PR 02-MAR-2001; 2001US-00798586.  
 XX  
 PA (MORRIS D W.  
 PA (ENGELHARD E K.  
 XX  
 PI Morris DW, Engelhard EK;  
 XX  
 DR WPI; 2004-328562/30.  
 XX  
 PT New carcinoma associated gene or protein, useful for preparing a  
 PT composition for diagnosing or treating carcinoma e.g., leukemia or  
 PT lymphoma.  
 XX  
 PS Claim 1; SEQ ID NO 148; 29pp; English.  
 XX  
 CC The invention relates to new recombinant nucleic acids. The invention  
 CC also relates to a host cell comprising a recombinant nucleic acid or  
 CC expression vector, an expression vector comprising a recombinant nucleic  
 CC acid, a recombinant protein, a method of screening for drug candidates, a

CC method of screening for a bioactive agent capable of binding to a  
 CC carcinoma associated protein (CAP) encoded by a nucleotide sequence, a  
 CC method of screening for a bioactive agent capable of modulating the  
 CC activity of a CAP, a method of evaluating the effect of a candidate  
 CC carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting  
 CC the activity of a CAP, a method of treating carcinomas, a method of  
 CC neutralising the effect of a CAP and a method of diagnosing carcinoma or  
 CC propensity to carcinoma. A method of evaluating the effect of a candidate  
 CC carcinoma drug comprises administering the drug to a patient, removing a  
 CC cell sample from the patient and determining alterations in the  
 CC expression or activation of a gene comprising the nucleotide sequence. A  
 CC method of diagnosing carcinoma comprises determining the expression of  
 CC one or more genes comprising the nucleic acid sequence in a first tissue  
 CC type of a first individual and comparing the expression of the gene from  
 CC a second normal tissue type from the first individual or a second  
 CC unaffected individual, where a difference in the expression indicates  
 CC that the first individual has carcinoma. A method of inhibiting the  
 CC activity of a CAP comprises binding an inhibitor to the CAP. Treating  
 CC carcinomas comprises administering to a patient an inhibitor of CAP.  
 CC Neutralising the effect of a CAP comprises contacting an agent specific  
 CC for the CAP. The polypeptide specifically binds to the protein encoded by  
 CC the nucleic acid. It comprises an antibody that specifically binds to the  
 CC protein encoded by the nucleic acid. The nucleic acids are useful for  
 CC preparing a composition for diagnosing or treating carcinoma e.g.,  
 CC leukaemia or lymphoma. This sequence represents a human carcinoma  
 CC associated (CA) nucleic acid of the invention. Note: The sequence data  
 CC for this patent did not form part of the printed specification but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html.  
 XX

SQ Sequence 22436 BP; 7692 A; 4872 C; 4296 G; 5576 T; 0 U; 0 Other;  
 Query Match 40.7%; Score 3049.8; DB 12; Length 22436;  
 Best Local Similarity 86.5%; Pred. No. 0;  
 Matches 3556; Conservative 142; Mismatches 320; Indels 93; Gaps 31;  
 QY 606 GATGGAAACGTTCCCGCAAGCAAAAACGCCCTTAAGACGTATTCTGGAATTTGGGA 665  
 Db 1581 GGTAGCAAAACATTCCTCTCAAGGCAAAAACACCCGAGATGTTCTGAGAAATCGGA 1640  
 QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAAGACATTAATTTCTTCGAGTCCGCCCTG 725  
 Db 1641 CCAATTTGACCTCAGACACTGAGAAAGAAATGACTTATATTTCTTCGAGTATTCCTG 1700  
 QY 726 GC-----ACTCCTGAGGGAAGTATAAATATAA 753  
 Db 1701 GCCAATATATCTTCTATGAGGGAAGAAACGTGGCTCTGAGGGAAGTATAAATATAA 1760  
 QY 754 CACCATTTACAGTAGACCTCTTTTGTAGAA---AAGCAAAATGGAGTGAAGTGCATA 810  
 Db 1761 CACCATTTACAGTAGACCTCTTTTGTAGAAAGAAAGCAATGGAGTGAAGGCGCATA 1820  
 QY 811 AGTACAAATTTCTTTTCAATTAAGAGACAACTCAAAATATGTAAAAAGTGTGATTTATG 870  
 Db 1821 TGTACAAATTTCTTTTCAATTAAGAGACAACTCGCTATATGTAAAAAGTGTGATTTATC 1880  
 QY 871 CCTACAGAGAGCCTTCAGAGTCTACCTCCCTATCCAGCAT--CCCCAGCTCTTCCCC 928  
 Db 1881 CCTACAGAGAGCCTTCAGAGTGTACCTTCCTACCCAGCATCCCCCTTCTTCCCC 1940  
 QY 929 AMYTAATAAGGACCCCTTCAACCAATGTGTCAAAAGGAGATAGACAAAAGGTATAA 988  
 Db 1941 AACTTAATAGGACCCCTTCAACCAATGTGTCAAAAGGAGATAGACAAAAGGTATAA 2000  
 QY 989 CAGTGAAACAAAGAGTGCATATTTCCCAATTTATGACCCCTTCCCAAGCAGTGGAGGAA 1048  
 Db 2001 CAATGAATCAAAAGAGTGCATATTTCTCGATTATGCCACT--CCAAAAGTGGAGGAG 2059  
 QY 1049 GAGAAATTCGGCCCGCAGAGTGCATGCTGCTTTTCTCCAGACTTAAGACAAATAA 1108  
 Db 2060 GAGAAATTCGGCCCGCAGAGTGCATGCTGCTTTTCTCTCAGTCTTAAAGCAAAATG 2119  
 QY 1109 AAACAGACTTAGGTAATTTCTCAGATAAATCTCAGATAAATCTGATGGCTATATTGRTGTGTTTACAAGGT 1168

Db 2120 AAATAGACCTAGGTAAATTTCTCAGCTAACCCCTGATGGCTATATTGATGTTTTACAAGGT 2179  
Qy 1169 TAGGACAAATCTTTGATCTGACATCGAGAGATATATATGTCTACTCTCTAAATCAGACACTA 1228  
Db 2180 TAGGACAAATCTTTGATCTGACATCGAGAGATATA-ATGTTACTCTTTAAATCAGACACTA 2238  
Qy 1229 ACCCCAAATAGAGAAGTGCACCAATAACTGACGCTGAGGTGTTGGCGATCTCTGGTAT 1288  
Db 2239 ACCCCAAATAGAGAAGTGCACCAATAACTGACGCTGAGGTGTTGGCGATCTCTGGTAT 2298  
Qy 1289 CTCAGTCAGGTCAATGGATPANGATGCAACAGAAAGAAAGANAAATGATCCCCACAGGC 1348  
Db 2299 CTCAGTCAGGTCCAT--CATAGGATGAAAAAG-AGGAAAGAGAACGATTCGCCACAGGC 2355  
Qy 1349 CAGCARGCAGTTCACAGCTASACCTCATTTGGGACACAGAAATCAGTAACATGGGAGA 1408  
Db 2356 CAGCAGGCAGTTCACAGCTGAGACCTCATAGGG--ACACAGAAATCAGAACAT--GGAGA 2411  
Qy 1409 TTGGTGTGCAGACATTTGCTAACTTTGCTGTCTASAAGGACTTAAGAAACCTASGAAGAA 1468  
Db 2412 ATGGTGCAGCAGACATTTGCTAACTTTGCTGTCTAGAAGACTTAAGAAACCTAGGAAG-A 2470  
Qy 1469 ARTCTAYGAATTAATCAATGATGTCACCATACACAGGGGAAGGAAGAAATCCTACT 1528  
Db 2471 AGCCTATGAATTAATCAATGATGTCAGTATAACA-AAGGGAAGGAAGAAATCCTACT 2529  
Qy 1529 GCCTTTCTGGAGAGACTAAGGAGGCAATGAGGAAGCGTCTCTGTCTACCTCACTCT 1588  
Db 2530 GCCTTTCTGGAAAGACTAAGGAGGCAATGAGGAGGAAACCTGTCTGTCACTGACTCT 2589  
Qy 1589 TCTGAAGGCCAATTAATCTTAAGCGTAAAGTTTATCACTCAGCTAGCTGCAGACATTAG- 1647  
Db 2590 TTGAAGGCCAATTAATCTTAAGGAAAGTTTATCACTCAGCTAGCTGCAGACATTAGA 2649  
Qy 1648 AAAAACTTCAAAAGTCTGCCCTAGCGCCGAGCAAAAACCTTAGAAACCTTATGAACCTTG 1707  
Db 2650 AAAAACTTCAAAAGTCAAGCTTAGCCCTGAGGACAAAACCTTAGAAACCTTACTGAACTTG 2709  
Qy 1708 GCAACVTGG--TTTTTATAATAGATCAGAGAGCAGCGGAAACAGGACAAACGGGA 1766  
Db 2710 GCAACCTCGTTTTTTTATAATAGATCAGAGAGCAGCGGAAATGGGACAAACGGGA 2769  
Qy 1767 TTAATAAAGGCCACCGTTTTAGTCATGACCCCTCAGGCAAGTGACCTTTGGAGCTCTG 1826  
Db 2770 -TAAAAAAGGCCACCGTTTTAGTTATAGCCCTCAGGCAAGTGACCTTTGGAGCTCTG 2828  
Qy 1827 GAAAAGGGAAGCTGGGCAAAATGCAATGCCCTAATAGGCTTTGCTTCCAGTGCCTCTAC 1886  
Db 2829 GAAAAGGGAATGCTGAGCAAAATGGGTGCTTAATAGGCTTGTCTCCAGTGCAGTCTAC 2888  
Qy 1887 AAGGACATTTAAAAAGATTATCAAGTACAAATAGCAAAATAGCCACCCCTCGTCCATGCCGT 2948  
Qy 1947 TATTTCAAGGGAATCACTGGAAGGCCACTGCCCGGCAAGGCTTTTGGAGTCAG 2006  
Db 2949 TATATCAAGGGAATCACTGGAAGGCCAGGGCCCCGGGGGATGAAGGCTTCTGAGTCAG 3008  
Qy 2007 AAGCCACTAACAGATGATCCACGACGAGACTGAGGGTGCCTGGGGCAAGGCCATCCC 2066  
Db 3009 AAGCCACTAACAGATGATCCACGACGAGACTGAGGGTTCGGGGCAAGCCACAGCCC 3068  
Qy 2067 ATGCCATCACTCAGAGCCCTGGGTATGCTTGACCAATGAGGGCCAGGAAGT---T 2123  
Db 3069 ATGCCATCACTCAGAGCCCTGGGTATGCTTGACCTTTGAGGSCCAGGAGGTTAACT 3128  
Qy 2124 GTCTCTGACACTGGTGGGTCTTCTTAGTCTTACTCTTCTGTCCCGGACAACTGTCT 2183  
Db 3129 GTCTCTGACACTGGCAGCGCTTCTCAGTCTTACTGTCTGTCTCTGACAACTGTCT 3188  
Qy 2184 CCAGATCTCTACTATTTCTGAGGGGTCCNTAAGCAGGCGAGCTCAGTACTATCTTCTTCC 2243

Db 3189 CCAGATCTGTCACTATCC---GAGGGGTCTTAGGACAGCCAGTCACTAGATAC-TTCTCC 3244  
Qy 2244 CAGCCACTAAGTATGAACGGGAGCTTTATCTTTTTCATGCTTTTCAATTTATGCT 2303  
Db 3245 CAGGCACTAAGTTATG-ACTGGGAACCTTTACTCATTTTCATGCTTTTCAATTTATGCC 3303  
Qy 2304 TGAAGGCCCACTACTCTTTGTAGGGAGAGACATTTCTAGCAAAAGCAGGGGCATTTATACA 2363  
Db 3304 TGAAGGCCCACTCTCTTTGTAGTGAGAGAAATTTCTAGCAAAAGCAGGGCCATTTACACA 3363  
Qy 2364 CTTGAACATAGGAGAGAACACCCGTTTGTGTGTCCTCTGTTGAGGAAGGAATTAATC 2423  
Db 3364 CTTGAACATAGGAGAGAACACCCGTTTGTGTG-CCCCTGCTTTGAGGAAGGAATTAATC 3422  
Qy 2424 CTGAAGTCTGGCAACAGAGGACAAATATGACAGGCCAAAGATGCCCTCTCTGTTCAA 2483  
Db 3423 CTGAAGTCTGGCAACAGAGGACAGTATGATGAGC--AAAGATGCACTCTCTGTTCAA 3480  
Qy 2484 GTTAAACTTAAAGGATTTCCACTTTCTCTTCCCTTACCAAGGCAAGTACCCCTCAGACCCCAAG 2543  
Db 3481 GTTAAACTTAAAGGATTTCTGCTCTCTTCCCTTACCAAGGCAAGTACCCCTTAGACCCCAAG 3539  
Qy 2544 GCCCAACAGGATTTCCAAAGATTTTAAAGACTTTAAAGCCCAAGGCTTTAGTAAAAACA 2603  
Db 3540 GCCCAACAGGACTTCCAAAGATTTTAAAGACCGAAAGGCCCAAGGCTTAGTAAAAACA 3599  
Qy 2604 TGCATACTCCTCGAGTAAATTCGTAGTGAATGAGGAGGACAGAAACCCAGTGGACA 2663  
Db 3600 TGCAATAGCCCTTCCAATATTCC-----AATTTTAGGAGTACAGAAACCCAGTGGACA 3652  
Qy 2664 GTGAGGGTGTAGTGCAGATCTCAGGATTAATCAATGAGGCGCTTGTCTTTTATACCACA 2723  
Db 3653 GTGGA-GGTTAGTGCAGATCTCAGGATTAATCGAT-TAGGCTGTGTTCTCTATATCCA 3710  
Qy 2724 GCTGTACTAGCCCTTATCTGTGMYTTTCCAAATPACAGAGGAAGAGAGTGGTTTACA 2783  
Db 3711 GCTGTACTAACCTTATACCTGCTTTTCCCAATACCAAGGAAGCAGAGTGGTTTACA 3770  
Qy 2784 STCTGGAACCTTMAAGATGCTTCTGTGATCCTCTGTACATCTCTGACTCTCAATTTCTTG 2843  
Db 3771 GTCTGGAACCTTAAAGGATGCTTTTCTGTGATCCTCTGTACATCTCTCAATTTTATA 3830  
Qy 2844 TTTGCTTTGAAGATACCTTCAAAACCAACATCTCAACTCAGCTGACCTTTTACCCCAA 2903  
Db 3831 TTTGCTTTGAAGATCTTCCGAAACCAATGCTCAACTCAGCTGACCTTTTACCCCAAG 3890  
Qy 2904 GGGTTCAAGGATAGYCCCCATCTATTTGGCCAGGATTTAGCCCAAGACTTGAGYCAATVM 2963  
Db 3891 GGGTTCAAGATAGCCCCCATCTATTTGGCCAGGCTTAGCCCAAGACTTGAGCCCAATTC 3950  
Qy 2964 TCATACCTGGACACTTGTCTCTGTGATGATGATTTACTTTTGTGCGCCCTTTAC 3023  
Db 3951 TCATACCTGGACACTTGTCTCTGTGATGATGATTTACTTTTGTGCGCCCAATTC 4010  
Qy 3024 GAAACCTTTGCGCATCAAGCCACCAAGCTCTTTMAATTTTCTCGYACCTGTGGCTAC 3083  
Db 4011 GAAGCTTTGCGCATCAAGCCACCAAGCACTCTTAAATTTCTCGCTACCTGTGGCTAC 4070  
Qy 3084 AWGGTTTCCAAACSAARAGCTCARTCTGTCTCAGACAGGTTTAAATACTAGGCTTAARA 3143  
Db 4071 AA-GATTTCCAAACCAAGGCTCAGCTCTGTCTCAGAGAGTTTAAATACTTAGGGCTAAAA 4129  
Qy 3144 TTATCCAAAGGACACAGGCCCTCAGTGAGGAARYATCCAGCTTATCTAGCTTATCTCT 3203  
Db 4130 TTATCCAAAGGATACAGGGCCCTCAGTGGGGAACATATGCAGCTTATCTAGCTTGTCTCT 4189  
Qy 3204 CATCYCAAAACCTTAAAGCAACTAGGRRRTTCTTTGGCTAAAYAGGVTTCCTGCCAAMA 3263  
Db 4190 CATCCAAACCTT-AAGCAACTAGAGGGTTCTTTGGCTAACAGGTTTCTGCCAATA 4248  
Qy 3264 TGGATTTCCCAAGGTGTCRAAATAGCCAGGYCATTAATACASTAATTAAGGAAACTCA 3323  
Db 4249 TGGATT-CCCAAGGTATGGCAAAATAGCCAGATCAATTATATACACTAAATTAAGGAAACTCA 4307











---